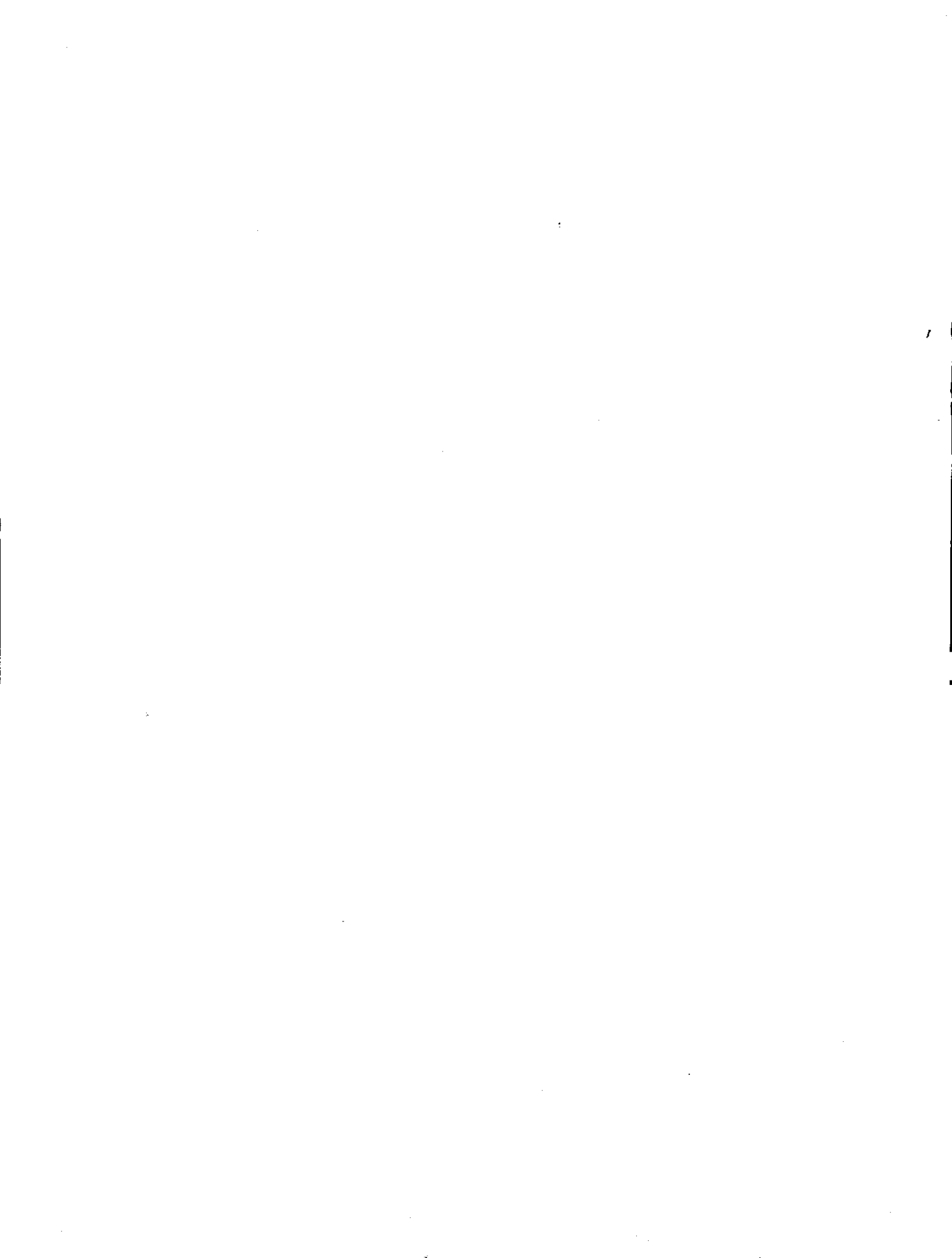


**LAS BASES DE DATOS TEXTUALES Y SU APLICACION
EN EL CAMPO LEGAL GUATEMALTECO**

BIBLIOTECA

UNIVERSIDAD DEL CAJON GUATEMALA



72
1104
Comp.
2825
1989

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ciencias y Humanidades

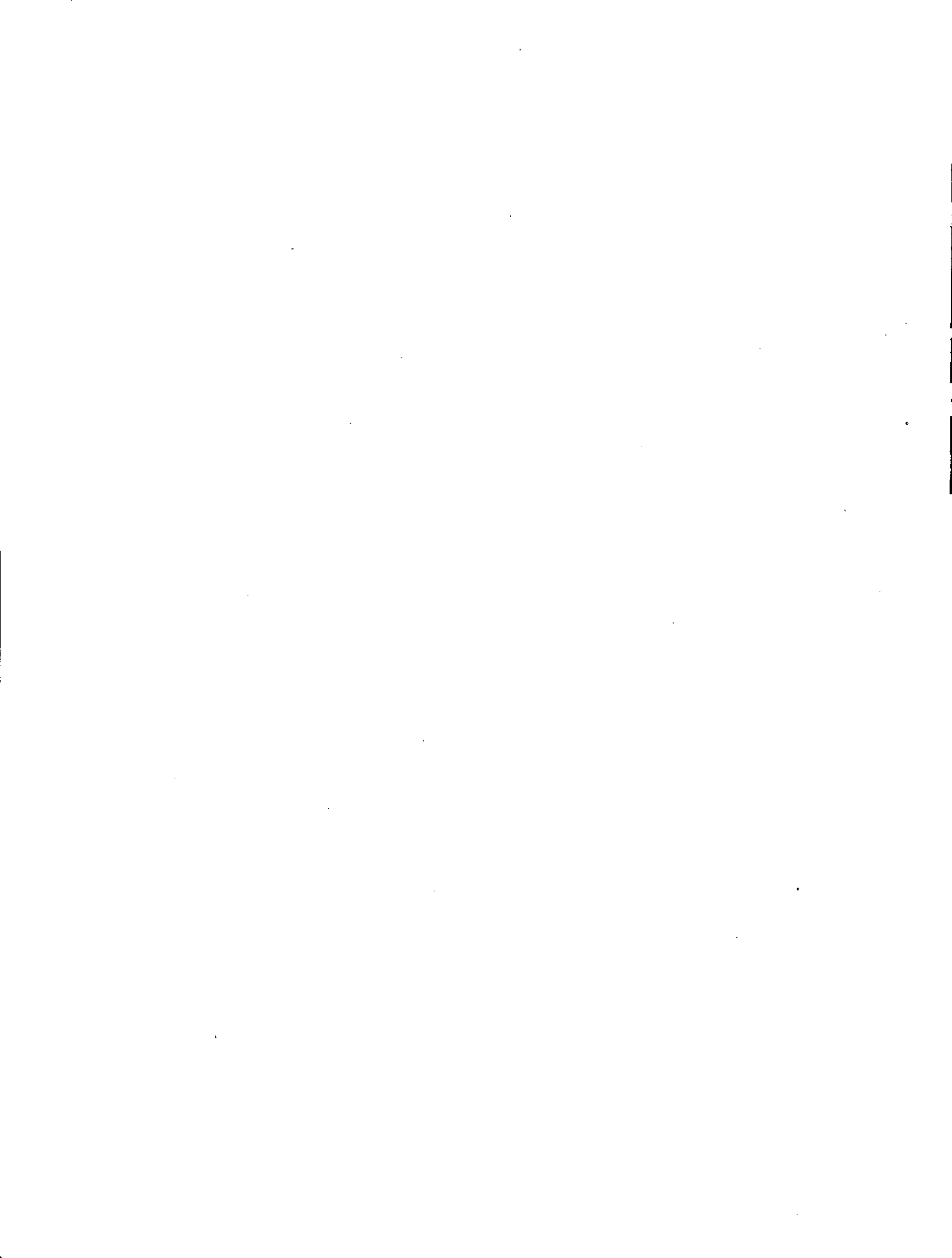
LAS BASES DE DATOS TEXTUALES Y SU APLICACION
EN EL CAMPO LEGAL GUATEMALTECO

FABIOLA SOTO MIDENA

Modelo de Trabajo Profesional presentado para optar al Grado
Académico de Ingeniera en Ciencias de la Computación
en el Grado de Licenciada

Guatemala

1989



Vo.Bo.:

(f) *O. cordón*
Licenciado Luis Octavio Cordón Girón
Asesor

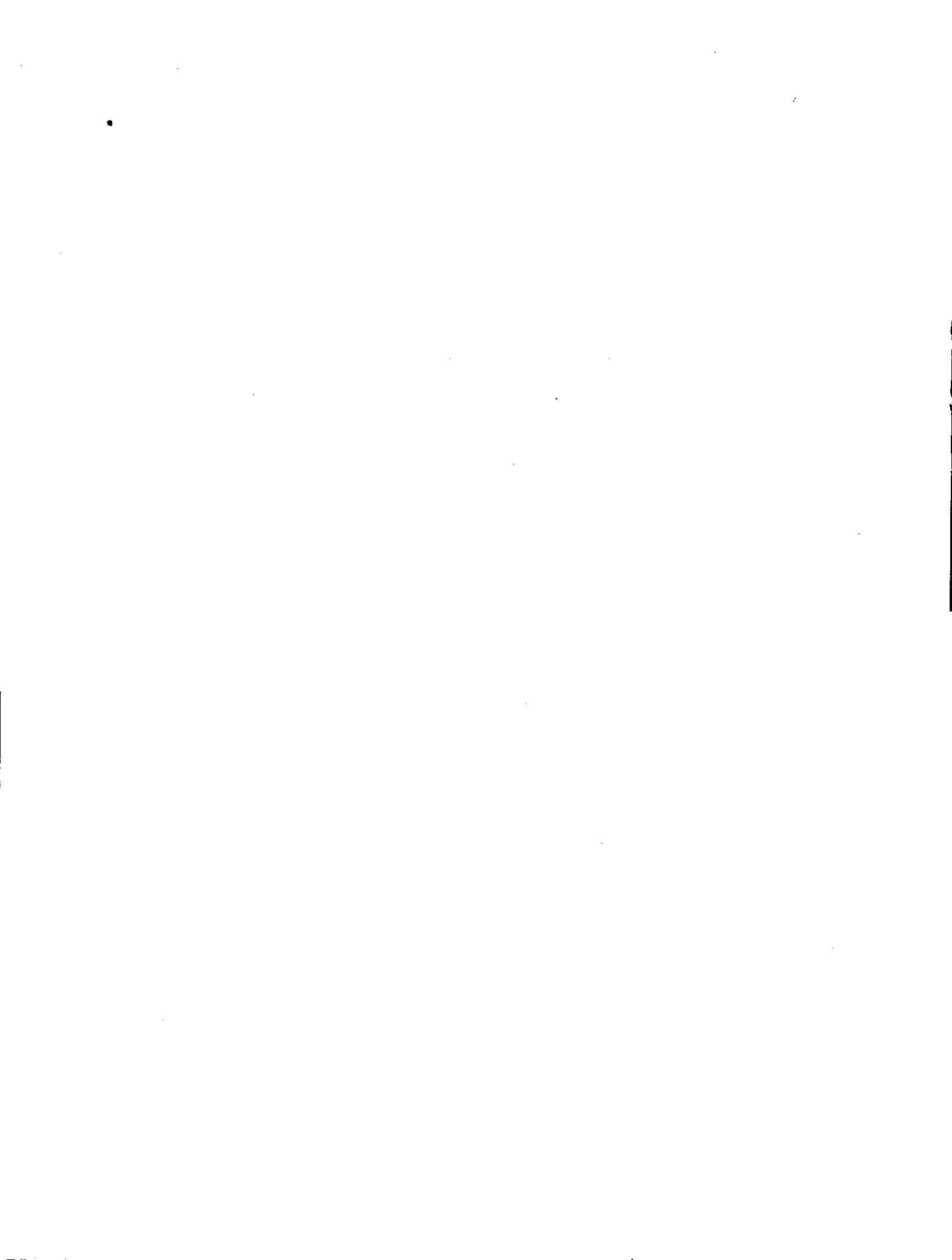
Comité del Trabajo Profesional:

(f) *Rodrigo Arias P.*
Licenciado Rodrigo Arias P.

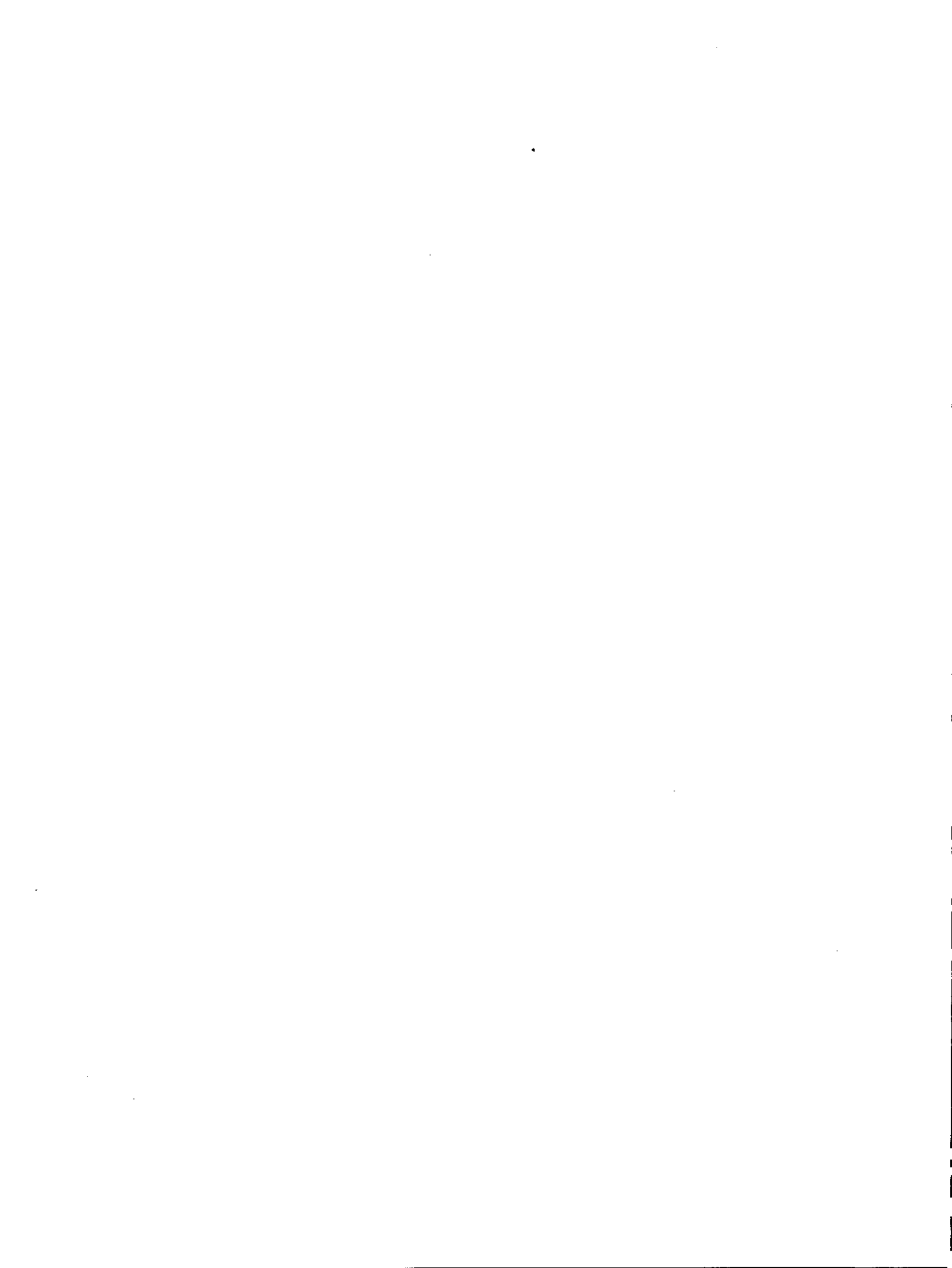
(f) *Luis R. Furlán*
Ingeniero Luis Furlán

(f) *O. cordón*
Licenciado Octavio Cordón

Fecha de aprobación: 6 de noviembre de 1989.



A todos los que hicieron posible este trabajo,
a mis padres, hermanas y amigos.



CONTENIDO

	Páginas
I. INTRODUCCION.	1
II. OBJETIVO.. . . .	3
III. FUNDAMENTOS TEORICOS DE LAS BASES DE DATOS TEXTUALES	5
A. Qué son las Bases de Datos ?	5
1. Arquitectura de un Sistema de Bases de Datos.	7
B. Tipos de Bases de Datos.	8
C. Diferencias entre una base de datos textual y una base de datos de formato fijo.	18
1. El Concepto de Formato Libre.	19
D. Problemática y teoría de las bases de datos textuales.	21
1. Historia de las bases de datos textuales.	21
2. Estado actual del desarrollo en las bases de datos textuales y los factores determinantes para la puesta en marcha de las mismas.	23
a. El Procesador de Esquemas o Modelo Jerárquico.	25
b. Modelo No Jerárquico.	28
3. Conceptos básicos para sistemas de recuperación de documentos.	31
a. El Concepto de Relevancia	32
b. El Concepto de Recolección	32
c. El Concepto de Precisión	33
d. El Concepto de Desacierto	34
4. Análisis de la eficiencia en sistemas de búsqueda para documentos.	35
a. Sistemas de Recuperación de textos completos.	35

	b.	El tamaño de los archivos y rendimiento de recuperación en sistemas de texto completo.	39
	5.	Importancia de la aplicación de bases de datos textuales y su lugar en el futuro del mundo actual.	43
IV.		UN AMBIENTE DE UN BUFETE DE ABOGACIA Y NOTARIADO.	45
	A.	Definición de conceptos de Informática Jurídica.	45
		1. La Informática Jurídica Documental.	46
		2. La Informática Jurídica de Gestión.	47
		3. La Informática Jurídica Decisional.	47
	B.	Descripción del ambiente de trabajo en un bufete profesional de abogacía y notariado.	48
		1. Control de Tiempo.	49
		2. Facturación automatizada.	50
		3. Sistema computarizado de inscripción, renovación y oposición de marcas, expresiones de propaganda y nombres comerciales.	51
		4. Protocolo del Notario.	51
		5. Sistema de Fechas de Vencimiento.	53
		6. Contabilidad y sistemas de cuentas por cobrar y por pagar.	54
		7. Reproducción de documentos.	55
		8. Investigación Legal asistida por computadora.	56
		9. Comunicaciones.	58
	C.	Tareas automatizables utilizando Bases de Datos Textuales.	59
V.		LA LEY DEL IMPUESTO SOBRE LA RENTA. UNA APLICACION ESPECIFICA.	61

A.	Descripción de la Metodología a usarse para desarrollar la aplicación.	62
1.	Diseño del paquete dayflo:	64
B.	Desarrollo de la aplicación "La Ley del Impuesto sobre la Renta".	72
VI.	EVALUACION DEL SISTEMA DESARROLLADO EN ESTE TRABAJO Y OTRAS APLICACIONES UTILIZANDO BASES DE DATOS TEXTUALES.. . . .	107
A.	Ventajas y beneficios de la utilización del sistema Dayflo	107
B.	Desventajas y limitaciones de la utilización del sistema Dayflo	108
C.	Evaluación de la utilidad del sistema Dayflo en un bufete de abogados y notarios	111
1.	Descripción y definición del material a analizar	112
2.	Desarrollo del análisis de utilidad del sistema Dayflo en el ambiente de un bufete de abogados y notarios	117
3.	Análisis y conclusiones de los resultados obtenidos en la investigación de utilidad de este sistema	128
VII.	SUGERENCIAS Y CONCLUSIONES.	133
VIII.	BIBLIOGRAFIA	139
	ANEXOS	
A.	Programas utilizados para desarrollar la aplicación	143
B.	Pantalla de definición de registros	149
C.	Elementos definidos en el diccionario	151
D.	Resultados de las consultas realizadas en el análisis de utilidad	153
E.	Hoja de solicitud de información	155
F.	Esquema de la arquitectura del sistema Dayflo	157



LISTA DE TABLAS Y GRAFICAS

Tabla		Página
6.1	Información tomada en cuenta en la evaluación. . .	113
6.2	Frecuencia de utilización por base de datos. . . .	118
6.3	Frecuencia de utilización por día	121
6.4	Frecuencia de utilización por semana	121

Gráfica		Página
6.1	Frecuencia de utilización por base de datos. . . .	119
6.2	Frecuencia de utilización por día	122
6.3	Frecuencia de utilización por semana	123
6.4	Precisión	125
6.5	Información adicional encontrada	127



I. INTRODUCCION

En la actualidad existe un gran desconocimiento de lo que son las bases de datos textuales y la aplicación que pueden tener en nuestro medio. Esto fue lo que me motivó a escribir y presentar este trabajo como proyecto de graduación.

Es pues, el fin de este trabajo, contribuir de alguna forma al progreso del país, dando a conocer el poder y utilidad que tienen estos sistemas hoy en día.

El ambiente que se escogió fue el campo legal guatemalteco, pero esto no excluye a los demás campos profesionales de poder aprovechar y utilizar este trabajo como una guía para aplicar estos sistemas de bases de datos textuales en sus tareas profesionales. Este trabajo presenta un ejemplo de las tantas aplicaciones que se le pueden dar a las bases de datos textuales.

Este trabajo consta de cuatro partes principales. En la primera parte se dan los fundamentos teóricos de las bases de datos textuales, pretendiendo presentar claramente toda la teoría existente en esta área.

La segunda parte abarca de una manera general lo que son los conceptos de la Informática Jurídica y el ambiente de trabajo en un bufete profesional de abogacía y notariado, subrayando todas aquellas actividades que de una u otra forma impliquen una

automatización usando bases de datos textuales.

En la tercera parte se presenta el desarrollo de una de las tantas aplicaciones del campo legal guatemalteco que se pueden desarrollar utilizando bases de datos textuales. La aplicación que se escogió desarrollar, en este trabajo, fue del campo de la Informática Jurídica Documental, o sea la consulta de la Ley del Impuesto sobre la Renta. Aquí se cubrieron aspectos como la determinación de disposiciones existentes en esta Ley en relación a una palabra o a un grupo de palabras, la generación de índices para la ubicación precisa del texto que contenga cierta palabra y la consulta selectiva de la Ley. La aplicación se desarrolló utilizando para ello una base de datos textual.

Como última parte se elaboró una evaluación de la utilidad que pueden tener estos sistemas de bases de datos textuales en un bufete profesional de abogados y notarios en Guatemala.

En este trabajo se pretende enfatizar la importancia que tiene un sistema de bases de datos textuales hoy en día y dar un ejemplo de la aplicación que puede tener en el campo legal guatemalteco, así como también exponer las variantes y características importantes que debe tener una base de datos textual. También se desarrolló una evaluación de la utilidad que tienen estos sistemas en un bufete de abogados y notarios.

II. OBJETIVO

El objetivo principal de este trabajo es proporcionar una herramienta útil para la consulta de grandes volúmenes de texto, por medio de una base de datos textual.

Para el efecto se dará un ejemplo sobre los beneficios que un sistema de este tipo puede tener para una persona que tenga que trabajar en sus tareas diarias con la consulta de grandes volúmenes de texto, tal es el caso de los abogados y notarios que dedican gran parte de su tiempo a este tipo de actividad.

En este trabajo se presentan los conceptos teóricos relativos a las bases de datos textuales y a la informática jurídica y se desarrolla una aplicación de la Ley del Impuesto sobre la Renta utilizando una base de datos textual para enfatizar la forma en que se pueden poner en práctica estos conceptos teóricos. Al final se presenta un análisis para evaluar si estos sistemas son o no útiles en la práctica.

III. FUNDAMENTOS TEORICOS DE LAS BASES DE DATOS TEXTUALES

En este capítulo se dan los fundamentos teóricos de las bases de datos textuales, con el propósito de presentar más claramente este trabajo.

A. Qué son las Bases de Datos ?

La humanidad ha tenido siempre la necesidad de recolectar datos, desde la edad de piedra hasta esta época del hombre moderno. Las computadoras han permitido que los usuarios puedan organizar y mantener los datos de una forma automática y eficiente. Las bases de datos han proporcionado también una forma de obtener mejor provecho de los datos que con los sistemas mecánicos o manuales ya existentes.

Una base de datos computarizada permite que la información sea organizada de muchas maneras, así como también sea buscada de formas diferentes en diversas aplicaciones.

Según los criterios que existen en el campo de la informática, una base de datos se puede definir como una colección de datos interrelacionados que se almacenan juntos sin redundancias innecesarias, para ser utilizados en múltiples aplicaciones. Los datos se almacenan de tal manera que sean independientes de los programas que los utilizan. Un procedimiento común y controlado debe existir para agregar, modificar y utilizar los datos en la base de datos. Los datos deben estructurarse de tal forma que provean un fundamento para

el desarrollo de futuras aplicaciones. Se dice que un sistema contiene una colección de Bases de Datos si estas son totalmente independientes en estructura.

Existen varias características importantes que hacen de las bases de datos sistemas de suma importancia y utilidad.

El control centralizado de los datos operacionales en las bases de datos es la ventaja más valiosa que poseen. Realmente, de esta centralización del control de los datos se deriva una serie de ventajas tales como el hecho de que la cantidad de redundancia de información se logra reducir al mínimo, simplificando los procesos de actualización de la información y la veracidad de la misma. Los problemas de inconsistencia de los datos se logran evitar hasta cierto punto. Los datos, al estar centralizados, pueden ser compartidos por distintos usuarios y aplicaciones. Se pueden establecer patrones generales para uniformar toda la información de la base de datos, así como también se pueden establecer restricciones de seguridad de la información. La integridad de los datos permite que no existan datos inconsistentes. Con la centralización de la información se logra balancear los conflictos de requerimientos del sistema, poniéndole prioridad a las tareas principales del sistema.

Otro punto que es bien importante es la independencia de los datos. Este factor es más que todo un objetivo que deben

alcanzar las bases de datos y no una ventaja inherente a las mismas.

Existen dos grandes razones por las que resulta importante que las aplicaciones sean independientes de los datos:

- 1) Aplicaciones distintas pueden necesitar los mismos datos, pero de una forma diferente.
- 2) El administrador de la base de datos (DBA) puede querer cambiar la estructura de almacenamiento o su estrategia de acceso, o ambas, según los requerimientos del sistema, sin tener que alterar todas las aplicaciones existentes, o viceversa.

En lo que a usuarios de las bases de datos se refiere, podemos considerar tres tipos: el programador de aplicaciones, el usuario final u operador de aplicaciones y el administrador de la base de datos (siglas en inglés DBA, Data Base Administrator).

1. Arquitectura de un Sistema de Bases de Datos.

Cada sistema de bases de datos tiene una arquitectura definida, pudiendo ésta ser diferente en cada caso. En general, esta arquitectura se puede dividir en tres niveles: el interno, el conceptual y el externo. El nivel interno es el que está más cercano al almacenamiento físico de los datos, y es en el que se determina la forma física como se almacenarán los datos. El nivel externo es el que está más

cercano al usuario, o sea que se encarga de la forma en que el usuario ve los datos en sus aplicaciones. El tercer nivel, el nivel conceptual, es el que sirve de conexión entre los dos anteriores.

Toda base de datos tiene también un sistema manejador de la bases de datos (siglas en inglés DBMS, Database Management System), que es el software que maneja todos los accesos a la base de datos. Este sistema manejador de la base de datos actúa como un programa que controla el flujo y operación de todo el sistema de la base de datos. Es también el responsable de aplicar las verificaciones de autorizaciones y los procesos de validación en la base de datos, así como también la comunicación y coordinación entre los niveles de la arquitectura de la base de datos.

B. Tipos de Bases de Datos.

En la actualidad, para clasificar a las bases de datos existen varios criterios. El que se ha usado desde el surgimiento de las bases de datos es el relativo al tipo de relaciones o asociaciones entre las entidades de la base de datos.

El corazón de cualquier sistema de base de datos es el modelo de los datos conocido también como el nivel conceptual, como se mencionó en la descripción de la arquitectura de las

bases de datos. La forma que este modelo de datos y su respectivo lenguaje de datos pueda tomar es la manera más conveniente para clasificar los sistemas de bases de datos. Las tres categorías más populares son:

- 1) el modelo RELACIONAL;
- 2) el modelo JERARQUICO; y
- 3) el modelo de RED.

Todo sistema de bases de datos debe poder representar dos tipos de objetos fundamentales, las "entidades" y las "asociaciones". Cabe notar que no existe una diferencia sustancial entre ambas, ya que una asociación es un tipo especial de entidad. Los tres modelos (relacional, jerárquico y de red), arriba mencionados, se diferencian por la forma en que permiten al usuario ver y manipular las asociaciones.

1) El Modelo RELACIONAL.

En el modelo relacional los datos se representan por "tablas" formadas de filas y columnas. Estas tablas se parecen a un archivo secuencial tradicional en donde las filas de la tabla corresponden a los registros del archivo y las columnas corresponden a los campos de los registros, sin embargo existen varias diferencias significativas.

En el modelo relacional, las asociaciones se representan igual que las entidades, o sea como tuplas en relaciones.

Cada tabla es un caso especial de la construcción que se conoce en la matemática como "relación". El término relación tiene una definición mucho más precisa que los términos tradicionales usados en el procesamiento de datos como lo son "archivo" o "tabla". El término "tabla" se utiliza como sinónimo de "relación" cuando se habla del modelo relacional.

El modelo relacional se basa en el hecho de que los archivos que cumplan con ciertas restricciones pueden ser considerados como relaciones matemáticas. Como consecuencia de esto, la teoría elemental de relaciones puede aplicarse a varios problemas prácticos en el manejo de estos datos.

En la mayoría de la literatura relacional se hace referencia a las tablas como relaciones. A las filas de estas tablas se les llama generalmente "tuplas", debido nuevamente a que este término tiene una definición más precisa que la "fila" o el "registro". Sin embargo, los términos tupla y fila son los más usados. A las columnas se les llama "atributos", pero estos dos términos (columna y atributo) son intercambiables.

En la figura 1.1 se muestra un ejemplo relacional, en donde cada tabla es una relación, las filas son tuplas y las columnas los atributos.

Figura 1.1

Ejemplo del modelo relacional

Empleados			Servicios Prestados			
NN	Nombre	Apellido	NN	Actividad	Tiempo	NC
3	Juan	Pérez	1	Visita	3.3	3
4	Antonio	Juárez	3	Charla	1.0	3
1	Valeria	Fuentes	1	Lectura	0.5	4
			1	Análisis	1.0	5
			4	Visita	2.5	2
			3	Junta	1.1	1
			1	Estudio	1.5	4

Clientes		
NC	Nombre	País
1	Casa T	Guatemala
2	AFELSA	El Salvador
3	PAN OK	Guatemala
4	COGRISA	Honduras
5	PEREZ	Guatemala
6	PRISTA	México
7	GILSA	Honduras
8	AVISA	Guatemala

Otro concepto que se maneja en la teoría relacional es el "dominio". El "dominio" es una colección de valores de la cual se derivan los valores reales que aparecen en una columna dada.

Una característica crucial de la estructura de datos relacional es que las asociaciones entre las tuplas son representadas únicamente por valores de datos en las columnas extraídos de un dominio común.

Otra característica del modelo relacional es que toda la información en la base de datos, o sea, las entidades y las asociaciones, se representa de una única manera uniforme, o

sea, en forma de tablas. Esta característica no es compartida por el modelo jerárquico o el de red.

La uniformidad de esta representación de datos implica una uniformidad correspondiente en el conjunto de operadores. Como la información es representada de una única forma, se necesita un único operador por cada una de las funciones básicas que se quieran ejecutar.

El sublenguaje de datos en este modelo relacional se basa en el "álgebra relacional". Los operadores SELECT, JOIN, PROJECT, etc. constituyen esta álgebra relacional. Cada operación del álgebra relacional toma una o dos relaciones como su(s) operando(s) y produce una nueva relación como su resultado.

2) El Modelo JERARQUICO.

En este modelo los datos se representan como una estructura de árbol. En este tipo de base de datos el usuario visualiza varios árboles individuales u ocurrencias jerárquicas.

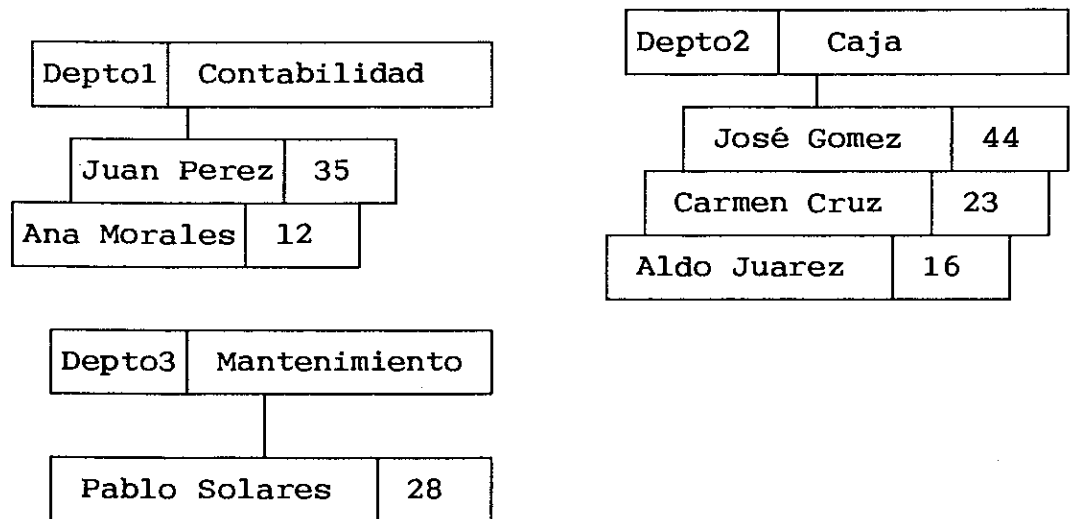
Obviamente, un archivo de este modelo jerárquico es un objeto mucho más complejo que las tablas del modelo relacional. En primer lugar, contiene varios tipos de registros y también "conexiones" vinculando las ocurrencias

de estos registros.

La figura 1.2 presenta un esquema del modelo jerárquico para bases de datos, en donde se presentan los empleados pertenecientes a un departamento dado de la empresa.

Figura 1.2

Ejemplo del modelo jerárquico



Para la vista jerárquica de los datos es fundamental el hecho de que cualquier ocurrencia de registro toma su significado completo sólo cuando es visto dentro del contexto general de la base de datos, o sea, ninguna ocurrencia de registro puede existir sin su vinculación preestablecida dentro de la estructura de árbol de este modelo.

En el sublenguaje de datos se utilizan operandos adicionales para especificar el contexto de los datos y así

identificar la conexión de la ocurrencia en cuestión.

Si se realizan dos preguntas sobre el modelo jerárquico, en donde ambas son simétricas, los procedimientos para resolver estas preguntas pierden la simetría, mientras que en el modelo relacional esta simetría de las preguntas se conserva. La pérdida de simetría es una consecuencia de este modelo, el cual es en sí asimétrico. En este modelo las entidades son tratadas como antecesores o sucesores, dependiendo de las asociaciones predefinidas. Esta asimetría es una gran desventaja en este modelo jerárquico, ya que genera complicaciones innecesarias para el usuario. El usuario se ve forzado a dedicar tiempo y esfuerzos en resolver problemas que son provocados por el modelo y no son intrínsecos a las preguntas realizadas. Esta situación se torna más difícil mientras se introducen más tipos de registro a la estructura y la jerarquía se vuelve más compleja. Este es un punto muy importante, ya que los programas se vuelven más complicados de lo necesario, provocando que escribir los programas, revisarlos y mantenerlos consuma más tiempo de programación que si el modelo no generara estas complicaciones.

Por otro lado, las jerarquías son una manera natural de representar estructuras del mundo real verdaderamente jerárquicas, donde se dan únicamente relaciones del tipo de

uno-a-varios. Las relaciones de varios-a-varios no son un ejemplo de una estructura jerárquica. En una situación jerárquica genuina también se da el problema de asimetría en la recuperación de registros. Estas situaciones jerárquicas genuinas tienden a generar con el tiempo situaciones de varios-a-varios más complejas.

En cuanto a las operaciones de almacenamiento, se puede decir que el modelo jerárquico posee ciertas propiedades indeseadas. Se generan anomalías innecesarias en las conexiones con las operaciones básicas de almacenamiento (insertar, borrar, actualizar).

3) El Modelo de RED.

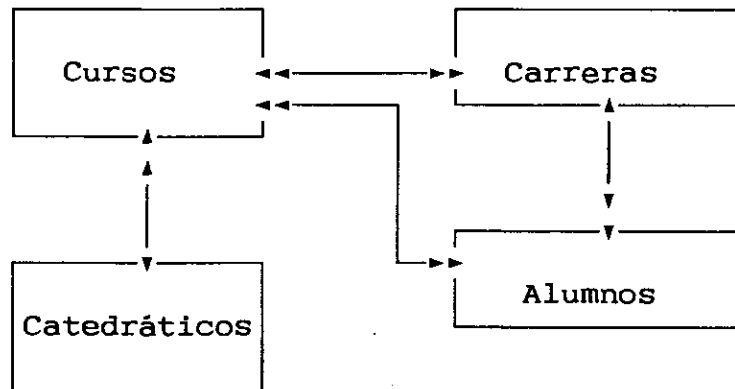
En este modelo, al igual que en el jerárquico, los datos se representan por "registros" y "conexiones". Una red es una estructura más general que la jerárquica, ya que una ocurrencia de registro puede tener cualquier número de superiores inmediatos, así como también cualquier número de dependientes inmediatos. En el modelo jerárquico los registros están limitados a un máximo de un superior inmediato. Esto significa que el modelo de red permite un modelo directo de la correspondencia de varios-a-varios.

En la figura 1.3 se muestra la estructura de una base de datos tipo red, en donde «←-----» es una relación de varios a

varios y $\leftarrow\text{-----}\rightarrow$ es una relación de uno a varios. El ejemplo muestra una organización educativa. Cada ocurrencia de los registros en este ejemplo tendría una o varias cadenas de punteros haciendo las conexiones entre las distintas entidades.

Figura 1.3

Ejemplo del modelo de red



Además de los tipos de registro que representan a los datos en la base de datos, existe otro tipo de registro que se puede denominar "conector". Una ocurrencia de este conector representa la asociación entre los otros registros de la base de datos y contiene datos que describen esta asociación. Todas las ocurrencias de conectores de una ocurrencia de un registro en la base de datos se representan por una cadena que se inicia en esta ocurrencia del registro de datos y regresa a esta misma. Estas cadenas pueden representarse físicamente por cadenas de punteros o por cualquier otro método equivalente.

Las jerarquías se pueden representar fácilmente en un sistema de red. El modelo de red es más simétrico que el jerárquico.

La principal desventaja del modelo de red es su complejidad, tanto en el modelo en si como en el sublenguaje de datos asociado. La fuente de esta complejidad está en el rango de construcciones relacionadas con la información que soporta el modelo de datos. En general, mientras más construcciones de este tipo existan, más operadores se necesitarán para manejarlas y, consecuentemente, el sublenguaje de datos se vuelve más complicado. El rango de estructuras de datos permitidos refleja el hecho de que el modelo de red es muy similar a una estructura de datos.

En el modelo jerárquico y en el de red, algunas asociaciones se representan por medio de conexiones. Básicamente estas conexiones son capaces de representar asociaciones de uno a varios; la diferencia entre el modelo de red y el jerárquico es que en el de red se puede combinar para formar asociaciones más complejas de muchos a muchos, mientras que en el jerárquico esto no es posible. Otra diferencia entre el modelo jerárquico y el de red es que en el de red las asociaciones tienen nombres concretos y en el modelo jerárquico estas conexiones permanecen anónimas.

C. Diferencias entre una base de datos textual y una base de datos de formato fijo.

Las bases de datos textuales o de formato libre se pueden considerar como un híbrido entre los procesadores de palabras convencionales y los sistemas tradicionales manejadores de bases de datos. Una base de datos de formato libre o textual puede usar cualquiera de los modelos de base de datos descritos anteriormente, pero su característica esencial es permitir manejar información de formato libre.

El enorme crecimiento de computadoras personales en los últimos años ha provocado una gran demanda de sistemas de bases de datos poderosos para computadoras pequeñas. Desafortunadamente hasta hace un par de años todos los sistemas disponibles usaban bases de datos de formato fijo, derivados de sistemas de cómputo grandes en los cuales este tipo de base de datos resulta adecuada. Sin embargo, este tipo de base de datos no siempre resulta adecuada como una herramienta personal de trabajo.

Existen dos razones de peso por las que se hace esta objeción. La primera es que a pesar de que uno quiera controlar una serie de elementos que tienen algún factor en común, en realidad dos registros no son siempre exactamente iguales. Alguien puede tener más de un número de teléfono, o una dirección en el extranjero, o un título inesperadamente largo. Es entonces que uno se encuentra con el problema de adaptar estos registros fuera de lo común en un formato que no permite excepciones.

La segunda razón es que no es natural esperar que todos los tipos diferentes de información que uno quiera registrar vayan a ajustarse perfectamente en categorías predefinidas.

1. El Concepto de Formato Libre.

Cualquier ambiente del mundo real que involucre individuos requiere mucha más flexibilidad que la permitida en el concepto de formato fijo. Una solución es mantener cada registro como único, pero conteniendo cualquier combinación de elementos comunes. El concepto de formato libre permite ingresar la información sin un objetivo predefinido y utilizarla posteriormente de alguna manera. Si se desea que la base de datos tenga estas posibilidades y características, el concepto de esquema predefinido o de formato fijo para la base de datos resulta inadecuado.

El concepto de Formato Libre en las bases de datos permite al usuario ingresar información en cualquier formato, ya sean texto, tablas, números, gráficas, etc., y recuperarla usando palabras claves presentes en la información ingresada. En estas bases de datos, generalmente la información carece de una estructura definida y se puede consultar con mucha facilidad. Con estas bases de datos uno puede concebir las computadoras ya no sólo como reproductoras de texto, sino que ahora los procesos de archivar y buscar documentos se reducen a rutinas rápidas y completas;

convirtiéndose la búsqueda de información en una actividad rápida y sencilla que permite que la creatividad y el intelecto del usuario sean utilizados al máximo.

La mayor parte de la información procesada hoy en día por los individuos es textual en su naturaleza, si se utilizan campos de formato fijo y números fijos de valores por campo se genera una situación incómodamente limitante. Adicionalmente, la naturaleza textual de los datos sugiere que la manera natural de crearlos y editarlos sea con un editor de texto.

Para que el sistema sea consistente, la organización, el orden y el método de consulta de los textos debe permitir que el usuario pueda realizar estas tareas con facilidad.

En las bases de datos de formato fijo los datos que la conforman están restringidos a un tipo de datos y formato predefinido donde no se permiten excepciones.

Dependiendo de la información que se pretenda manejar, así se debe escoger entre una base de datos de formato fijo o una de formato libre. La tendencia es utilizar las bases de datos textuales para aplicaciones que necesiten trabajar con mucho material escrito. Estos sistemas de formato libre se adaptan mucho mejor al ambiente de trabajo de cualquier

persona, debido a que simulan, hasta cierto punto, el ambiente de papeles en el que nos desenvolvemos diariamente. Lo cierto es que desde la antigüedad, el hombre ha utilizado el medio escrito y gráfico para plasmar sus conocimientos y experiencias, y hoy en día una base de datos textual o de formato libre proporciona al hombre actual una herramienta poderosa para utilizar este material escrito.

D. Problemática y teoría de las bases de datos textuales.

1. Historia de las bases de datos textuales.

Las bases de datos textuales son un concepto que pertenece al amplio mundo de la informática. Es por esto que cabe mencionar los rasgos históricos determinantes en el surgimiento de esta área.

Los factores que influyeron en el surgimiento de la informática fueron, entre otros, la crisis económica sufrida en los años 30 que desembocó en la segunda guerra mundial. La crisis económica creó un campo extenso de inventiva y desarrollo para poder mantener la producción, bajando los costos al máximo y delegando los procesos repetitivos a sistemas mecánicos más eficientes. En la milicia, las estrategias de guerra también requerían de cálculos exactos y complicados, que con el apoyo económico de las partes interesadas fue posible desarrollar máquinas de cálculo adecuadas, las cuales son los antepasados más semejantes a

las computadoras actuales.

Se puede decir que no fue hasta los años 70 en que se empezaron a desarrollar los primeros sistemas de bases de datos textuales. A inicio de los 80 salen las primeras bases de datos textuales al mercado. Antes de esto sólo existían o paquetes de bases de datos convencionales de formato fijo o programas procesadores de palabras, pero no la combinación de ambos. De 1984 a la fecha estos paquetes se han ido desarrollando y se han convertido en sistemas más completos de información en donde los datos son formados de textos, gráficas, números, así como también de otras combinaciones de estos, utilizando como base el concepto de formato libre, donde los datos del sistema no tienen restricciones en forma y tamaño. Se les ha puesto el nombre de "hipertextos" (Hypertext, en inglés), y no es sino hasta el segundo semestre de 1988 que pasan de la investigación en los laboratorios de Estados Unidos al mercado de paquetes de software. Las características básicas de un sistema de hipertexto es que provee acceso por medio de un simple teclazo, en casi todos los casos, a textos relacionados sin la necesidad de hacer búsquedas; establece relaciones fijas entre textos y posee indicadores visuales de que un texto o elemento de información está relacionado con otros.

2. Estado actual del desarrollo en las bases de datos textuales y los factores determinantes para la puesta en marcha de las mismas.

El hecho de integrar las funciones de las bases de datos y de edición simplifica el proceso creativo.

Resulta muy difícil hacer un reporte de una gran cantidad de datos textuales. Si uno utiliza un programa tradicional de procesamiento de palabras para almacenar elementos individuales mientras va desarrollando el material, uno termina con un gran grupo de archivos separados para colocarlos en uno sólo. Dependiendo del número de archivos así será la dificultad para leer cada documento e integrarlo en uno solo, así como también saber con exactitud qué contiene cada archivo.

Otra solución sería controlar la información de los documentos por medio de un programa de bases de datos. Cada pedazo de texto se etiqueta con un par de palabras claves y luego en las búsquedas se sabe cuáles elementos formarán el nuevo texto. El problema está en que estos programas de bases de datos no están diseñados para manipular textos completos como en un procesador de palabras. Entonces, el proceso para formar el nuevo documento debe realizarse casi manualmente con base en la información recuperada del sistema manejador de bases de datos.

Ninguna de estas dos soluciones es realmente buena y práctica. Un procesador de palabras puro carece del mecanismo de recuperación de información de las bases de datos, y un sistema de bases de datos carece de las facilidades de edición de un procesador de palabras.

En 1984 empezaron a surgir una serie de paquetes de software que trataron de resolver este dilema, tal como se explicó en la sección anterior. A estos productos se les conoce como "bases de datos textuales" o "procesadores de ideas". Estos programas son una combinación de editores de texto y segmentos de recuperación de bases de datos que permiten localizar rápidamente fragmentos de texto e incorporarlos en un documento más largo.

La integración de edición de texto y del manejador de textos, como un esfuerzo para proporcionar herramientas de software para intensificar la creatividad humana, ha resultado en programas ampliamente divergentes en sus guías filosóficas.

En el diseño de software una de las áreas más debatidas es que si el software debe conducirlo a uno a métodos más productivos de trabajo o debe dejarlo a uno reproducir de la manera más semejante sus hábitos de trabajo actuales, los cuales, seguramente, son muy desorganizados.

Los modelos que se presentan a continuación son genéricos, o sea que no son descripciones de programas específicos que existan actualmente en el mercado. Sin embargo varios productos pueden asemejarse bastante a estas descripciones. El mejor diseño para bases de datos textuales es difícil de determinar.

a. **El Procesador de Esquemas o Modelo Jerárquico.**

Los procesadores de esquemas son, ciertamente, una mejora a los procesadores de palabras convencionales, ya que permiten organizar el material textual como una especie de esquemas indentados. Uno tiene que construir el esqueleto de los niveles de tópicos y subtópicos y ya con esto puede realizar búsquedas sobre los esquemas de los textos completos guardados usando esta organización.

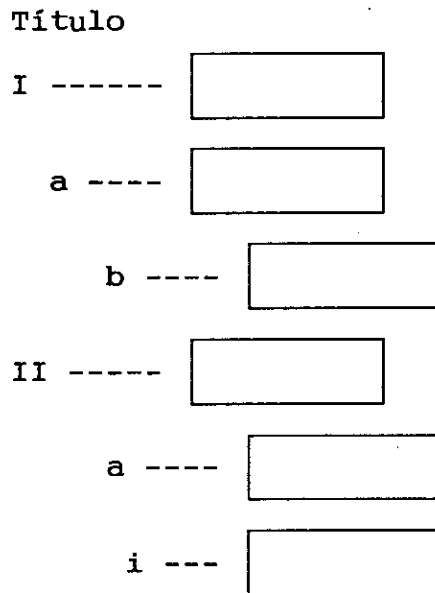
Si uno empieza a recolectar las notas o documentos de interés con un procesador de esquemas, uno puede mantener esta información bajo control. Debido a que la mayoría de los procesadores de esquemas pueden leer archivos externos como secciones dentro de los títulos de los esquemas, uno puede importar datos de otras fuentes, tales como procesadores de palabras tradicionales y manejadores de bases de datos.

Un ejemplo del modelo de esquemas es el que se

muestra en la figura 1.4.

Figura 1.4

Modelo de Esquema Indentado



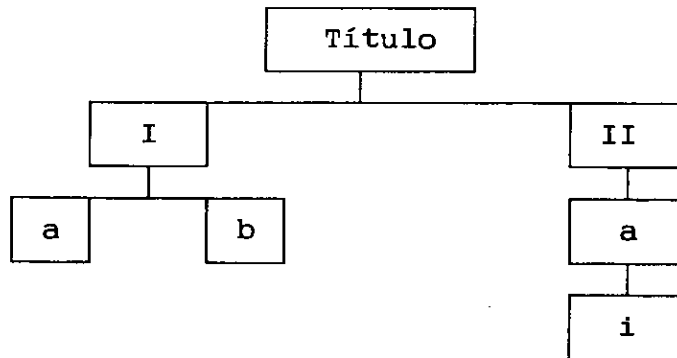
Este modelo es de estructura vertical en su forma más sencilla, pudiéndosele llamar modelo de esquemas indentado. En este modelo los textos tienen un lugar y un orden preestablecido, lo que crea una rigidez en el ingreso y organización de la información.

Si este modelo se invierte a una estructura horizontal de tal forma que parezca un árbol invertido. En este caso el orden cronológico de ingreso de los elementos textuales no es tan importante como en el esquema indentado, ya que se puede agregar y eliminar fácilmente cualquier elemento textual dentro de esta estructura de árbol invertido. Esta variación del modelo

de esquemas sigue siendo esencialmente jerárquico, no permitiendo flexibilidad en la estructura organizacional de los elementos textuales. A continuación se muestra la figura 1.5, la cual representa este modelo de esquema tipo árbol invertido.

Figura 1.5

Esquema tipo árbol invertido



De cualquier manera, esta rigidez en la estructura de ingreso de la información al sistema inhibe la espontaneidad. Un procesador de esquemas es, esencialmente, jerárquico en su estructura, en donde un elemento sigue a otro, y la organización puede llegar a ser un verdadero dolor de cabeza. La toma de notas al azar no es permitida en este sistema, y si uno no ha desarrollado una imagen clara del reporte final esperado se ve en problemas a la hora de adaptar un nuevo elemento al sistema existente.

A pesar de que estos procesadores de esquemas pueden

expandir y desmontar la vista de los textos dentro de un documento, ellos están limitados tanto en el tamaño del documento como en sus funciones sobre los textos, o sea, ellos no están diseñados para manejar grandes volúmenes de texto que requieran ordenamientos y reorganizaciones complicadas.

Este modelo de esquemas es una especie de base de datos, ya que organiza los bloques de texto llevando un control sobre los títulos, etiquetas o palabras claves del texto. Esta característica impone una estructura rígida sobre los textos y el proceso creativo, pero puede ser útil para mostrar una lista esquemática de todo el conjunto de textos.

Un procesador de esquemas es ideal para generar un reporte por vez, pero no es una buena herramienta para manejar un banco grande de datos.

b. Modelo No Jerárquico.

La forma más simple de un manejador de texto consiste en un editor con dos modos de operación, el modo "registro", y el modo "documento". En el modo "registro", el programa crea registros que son almacenados en la base de datos. Los programas más simples de este tipo usan registros cortos que son

identificados por un número limitado de palabras claves. En el modo "documento", el programa funciona como un procesador de palabras. Los comandos de edición de texto son iguales en los dos modos de operación, pero el modo "registro" tiene un conjunto adicional de comandos para archivar los registros individuales.

Existen dos formas principales para utilizar los datos textuales en este modelo no jerárquico. Uno es localizando los registros de texto por medio de palabras claves y leerlas al reporte para su edición, y la otra es localizando los textos por medio de una búsqueda completa en todo el contenido de estos textos.

La base de datos de "formato libre" es un modelo no jerárquico. A cualquier bloque de texto creado por el editor puede asignársele una o varias palabras claves que lo identifiquen y también puede almacenársele en la base de datos. Un registro puede ser llamado, ya sea utilizando una palabra clave o por medio de la búsqueda de una hilera de caracteres dentro de todo el texto, como se mencionó en el párrafo anterior. Obviamente el editor debe tener una especie de espacio de almacenamiento intermedio en caso de que una palabra clave dada o hilera de caracteres dada produzca una gran cantidad de fragmentos de texto. El editor puede

archivar individualmente cualquier bloque de texto que cree, y la base de datos puede hacer un índice sobre un grupo entero de bloques de texto a una sola palabra clave, de tal manera que un documento entero pueda ser construido de sus componentes. No es necesario diferenciar entre un registro y un documento; el sistema de formato libre es, esencialmente, carente de modos de operación.

El editor es el único ambiente visible. Mover y buscar datos en los alrededores de la base de datos puede realizarse por medio de comandos de ordenamiento y búsqueda dentro del editor de textos. La única excepción a esta regla es un programa que deja trabajar sobre varios archivos al mismo tiempo, en cuyo caso existe un estado intermedio en el cual se puede tener una visión amplia de los documentos disponibles.

El texto desarrollado con el editor es organizado con palabras claves y almacenado en la base de datos, la cual, generalmente, permite manejar grandes volúmenes de datos.

El mismo editor contiene comandos de búsqueda a la base de datos que permiten, en un momento dado, extraer todos los fragmentos de texto relacionados con el tema

cuestionado. Esto se hace ya sea buscando los textos por palabras claves o por hileras de caracteres dentro de los textos ingresados en la base de datos. El editor puede también almacenar texto en la base de datos para recuperarlo después. Este sistema es inherentemente no jerárquico.

La flexibilidad de este tipo de programa salta a la vista. Uno no se ve limitado a las restricciones del formato fijo donde la longitud de los registros es limitada. Debido a que los documentos pueden recuperarse completos en cada búsqueda como elementos individuales, la longitud de los mismos no es un punto determinante.

El hecho de asignar palabras claves que unan datos relacionados hace posible construir en un reporte materiales de diversos tipos, incluyendo bloques cortos de texto, documentos largos, memorándumes, cartas, y formas de registros. La base de datos puede crecer enormemente y de hecho, el programa se vuelve más útil mientras más datos existan; mientras más extenso sea el universo de datos a consultar mejor será la información que de él se pueda extraer.

3. Conceptos básicos para sistemas de recuperación de documentos.

Las medidas para evaluar los sistemas de recuperación de

documentos son recolección, precisión y desacierto. Junto a las definiciones de estas medidas se encuentra vinculado directamente el concepto de relevancia. En la práctica se sabe que la recolección está relacionada inversamente con la precisión, y que es difícil recuperar todo lo que sea relevante y rechazar lo irrelevante en las búsquedas.

a. El Concepto de Relevancia

Relevancia se puede definir como la relación entre una pregunta y un documento que resulte importante a la persona que esté realizando la investigación sobre el conjunto de textos en la base de datos. La relevancia es un concepto extremadamente complicado en las leyes. Cualquier estudio sobre recuperación de documentos necesita siempre una manera sencilla para determinar la relevancia.

En los estudios más aceptados, todas las decisiones de relevancia se dejan a criterio del usuario. Los documentos relevantes son aquellos que el usuario determina.

b. El Concepto de Recolección

La recolección es el porcentaje del número total de documentos relevantes en una base de datos que son recuperados por una búsqueda realizada. Por ejemplo, si

una colección de documentos contiene 20 documentos relevantes a la pregunta realizada, pero al hacer la búsqueda sólo 7 son recuperados, entonces se dice que el porcentaje de recolección fue de un 35%, ya que sólo el 35 % de todos los documentos importantes fueron recuperados. En algunos experimentos realizados, el valor promedio de recolección en los sistemas de recuperación de texto ha sido de un 20 por ciento.

En el campo legal el porcentaje de recolección debe optimizarse para que el sistema sea efectivo.

La recolección es obviamente difícil de medir. El número total de documentos relevantes encontrados por una búsqueda se puede determinar con facilidad, pero el número de documentos relevantes en una base de datos que no hayan sido encontrados por la búsqueda, casi nunca se sabe.

c. El Concepto de Precisión

Precisión es el porcentaje del total de documentos recuperados por una búsqueda que son relevantes a la pregunta. Por ejemplo, si una búsqueda recupera 10 documentos, de los cuales 7 son relevantes, entonces la precisión es de un 70%.

La precisión no es tan importante como la recolección en la investigación de leyes, pero es, sin embargo, una medida importante. El tiempo de un abogado es precioso, y el tiempo perdido en la lectura de documentos irrelevantes al caso puede tener repercusiones económicas al cliente. Al contrario de la recolección, la precisión es fácil de medir, para calcularla no se necesita saber la parte relevante que no fue recuperada en la búsqueda.

d. El Concepto de Desacuerdo

Desacuerdo es la proporción del número total de documentos irrelevantes de una colección recuperada por una búsqueda. A diferencia de la precisión y la recolección, lo que constituye un valor aceptable para el desacuerdo depende del tamaño de la colección de documentos. Un uno por ciento de desacuerdo en una colección con 400 documentos irrelevantes le darían al usuario sólo 4 documentos irrelevantes para leer, pero el uno por ciento de desacuerdo en una colección de 400,000 documentos irrelevantes inundarían al usuario con 4,000 documentos irrelevantes. Debido a que las bases de datos de investigación legal tienden a ser grandes, solo niveles extremadamente bajos de desacuerdo son aceptables en estos sistemas.

Como en la precisión, el desacierto es, relativamente, fácil de calcular para búsquedas individuales. El número de documentos irrelevantes recuperados, generalmente, se conoce. El número de documentos relevantes en una colección es una fracción trivial de la colección completa, y así el número total de documentos irrelevantes es poco más o menos igual en tamaño al de la colección, y este número, generalmente, se conoce.

4. Análisis de la eficiencia en sistemas de búsqueda para documentos.

La utilidad de un sistema de recuperación de documentos depende de su habilidad de proporcionar un sustituto adecuado a la construcción humana de índices para la información. Uno de los problemas más notorios en los sistemas automáticos de recuperación de documentos es que existe una correspondencia imperfecta entre las ideas y las palabras.

a. Sistemas de Recuperación de textos completos.

En la recuperación automática de textos completos no existe la construcción humana de índices a la que se acostumbra en sistemas manuales. La utilidad que pueda tener un sistema de este tipo depende del grado de sustitución que se logre a la construcción humana de índices. Las deficiencias de un sistema de este tipo

dependen exclusivamente de la inhabilidad de la máquina para diferenciar perfectamente entre un documento relevante a una pregunta y uno irrelevante.

Actualmente, estos sistemas pueden seleccionar los textos de la base de datos que contengan o se relacionen con las palabras proporcionadas por el usuario. El problema está en que existe una correspondencia imperfecta entre ideas y palabras. Los problemas específicos que surgen de estos sistemas se pueden clasificar de la siguiente forma: palabras sinónimas, palabras ambiguas, truncación de términos y expresiones complejas.

Palabras sinónimas.

El primer problema al que uno se enfrenta con el acoplamiento de palabras en el lenguaje natural es que existen varias palabras diferentes para referirse a la misma cosa.

Cuando uno hace la búsqueda debe evitarse que los elementos a buscar se puedan expresar en varias formas diferentes.

Es simplemente imposible anticipar todas las palabras que puedan ser utilizadas para referirse a una

situación o idea en el texto de una opinión. Si el usuario que realiza las búsquedas no conoce el lenguaje utilizado para describir los asuntos de interés, se corre el riesgo de que casos relevantes no sean encontrados y el factor de recolección sufra las consecuencias.

Tales términos casi siempre están disponibles en los diccionarios, la jerarquía de los términos ó pueden ser sugeridos por el usuario a la hora de hacer la búsqueda.

Palabras ambiguas.

Las palabras ambiguas crean el problema opuesto al de sinónimos. Así como una idea puede ser expresada de varias formas diferentes, una palabra puede, a su vez, representar ideas diferentes.

Los problemas de ambigüedad generan una ejecución pobre en precisión y desacierto. Si el problema de ambigüedad de una palabra es muy serio, la búsqueda de ese elemento será un fracaso.

Truncación de términos.

La truncación de términos consiste en utilizar términos truncados, o raíces de palabras, en lugar de los términos completos para la identificación de documentos en las preguntas. Por ejemplo al preguntar

por la raíz "cancela" se están contemplando los términos "cancelación", "cancelado", "cancelar", etc. Esto amplía el campo de acción más que si se utilizaran, únicamente, las palabras completas.

Expresiones complejas.

El problema más difícil con sistemas de texto completo es que varios de los significados de los textos de lenguaje natural están contenidos en relaciones complejas entre palabras y no en las palabras individuales en sí. A veces ni las relaciones lógicas más complicadas entre las palabras pueden servir para encontrar los textos que significan exactamente la idea buscada.

En los casos en que las palabras que expresan los elementos de una búsqueda son suficientemente distintivos, las propias palabras dan la información necesaria para que la computadora pueda distinguir entre los casos relevantes y los irrelevantes.

El problema existe ya que el vocabulario empleado en los textos de los sistemas de texto completo no tiene un control estricto. Estos sistemas son extremadamente analíticos. Con índices se identifican, únicamente, las partes de los documentos que son más significantes o

importantes, pero en los sistemas de texto completo todas las palabras son tratadas como elementos iguales. La ventaja es que se puede obtener al mínimo detalle cualquier fragmento de texto que cumpla con la condición de búsqueda, mientras que los sistemas que utilizan índices sólo recuperan los textos marcados, previamente, como de mayor importancia. La selección de cualquiera de estos dos sistemas, depende del tipo de sistema que se quiera poner en práctica.

b. El tamaño de los archivos y rendimiento de recuperación en sistemas de texto completo.

Uno de los factores que tiene un efecto dramático en el rendimiento de sistemas de texto completo es el tamaño de la colección de datos sobre la cual se realice la búsqueda; mientras más grande sea esta colección, menor será el rendimiento. Una herramienta que resulta útil para visualizar estos conceptos es la figura 1.6.

Figura 1.6

Situación generada por los sistemas de búsqueda de documentos

	Relevante	Irrelevante	
Recuperado	a	b	N1
No Recuperado	c	d	
	N2		N

En esta figura, N es el número total de documentos en la base de datos, $N1$ es el número de documentos recuperados en la búsqueda y $N2$ es el número de documentos relevantes a la pregunta de la búsqueda.

Estos N documentos de la base de datos se pueden, a su vez, dividir en cuatro clases cada vez que se realiza una búsqueda:

- tipo a documentos relevantes recuperados;
- tipo b documentos irrelevantes recuperados;
- tipo c documentos relevantes no recuperados; y
- tipo d documentos irrelevantes no recuperados.

El índice de "recolección" para una búsqueda se puede expresar, entonces, como:

$$\text{Recolección} = a/N2$$

y representa la fracción del número total de documentos relevantes que fueron recuperados por la búsqueda. El índice de "precisión" puede expresarse, similarmente, como:

$$\text{Precisión} = a/N1$$

y el índice de "desacierto" como:

$$\text{Desacierto} = b/(N-N2)$$

En los sistemas de recuperación de documentos se dan dos tipos de errores. Estos dos tipos de errores están representados en la figura anterior por las letras c y b. Los abogados están particularmente interesados en el error tipo c (documentos relevantes no recuperados), el cual contribuye a una mala "recolección". Los errores tipo b (documentos irrelevantes recuperados), contribuyen a una mala "precisión" y a un mal "desacierto". Estos errores tipo b, generalmente, no son de mucha trascendencia para los abogados, pero limitan la calidad de una búsqueda, tanto como el error tipo c.

Sólo un número pequeño de errores tipo b son tolerables en una búsqueda, ya que es indeseable tener que leer o analizar una serie de documentos recuperados que a la larga resultan irrelevantes a la búsqueda original. Los errores tipo b pueden eliminarse a costa de crear errores tipo c adicionales. Y como los errores tipo b están relacionados al tamaño de la colección de datos, mientras mayor la colección, mayor será el número de errores tipo c que se puedan esperar.

Blair y Maron en su artículo titulado en inglés "An Evaluation of Retrieval Effectiveness for a Full-text Document-retrieval System" publicado en la revista Communications of the ACM¹ examinan el sistema de

¹ "Communication of the ACM". Vol. 28, No. 3, Marzo 1985.

recuperación de textos STAIRS aplicado a una colección de 40,000 documentos de texto completo para contestar 40 preguntas diferentes hechas por el usuario. En STAIRS, las palabras normalmente se extraen del texto para identificar su contenido. Después de que cada palabra ha sido ensanchada usando truncación, el usuario puede proporcionarle su respectiva lista de sinónimos. Cuando estas listas han sido llenadas, cualquier búsqueda realizada sobre algún término se extiende automáticamente a la lista de sinónimos. El sistema STAIRS, también incluye una característica de clasificación, que recupera documentos en orden decreciente basada en el peso de las palabras, el cual se calcula agregando los pesos de los términos contenidos en cada documento recuperado.

A pesar de que muchas de las características del sistema STAIRS no son muy atractivas, STAIRS es, sin duda, uno de los sistemas de recuperación de texto más estándar que existe. En el experimento de Blair y Maron, se obtuvo un valor promedio de precisión de 75% y un valor promedio de recolección de 20%. Esto es, por cada una de las búsquedas realizadas, tres de cada cuatro documentos recuperados eran, de hecho, pertinentes a la pregunta del usuario, y, aproximadamente, un quinto del total de documentos relevantes presentes en la base de

datos fueron recuperados.

5. Importancia de la aplicación de bases de datos textuales y su lugar en el futuro del mundo actual.

La evidencia resultante de estudios que comparan sistemas de recuperación de texto manual y automático, no demuestran la conclusión de que el análisis intelectual del contenido produzca mejores resultados que sistemas automáticos semejantes.

Los sistemas de bases de datos textuales son una herramienta que permite que los usuarios dejen de pensar que las computadoras son, únicamente, una forma de reproducir papeles en un medio electrónico. Archivar y buscar documentos se reduce a rápidas rutinas; la investigación de fuentes de información se convierte en una tarea corta y fácil en el proceso creativo. Consecuentemente se puede emplear más tiempo en el trabajo creativo que en labores de búsqueda e investigación de documentos.

Las oficinas de abogados y notarios son un ejemplo clásico de una oficina que necesita un sistema computarizado de recuperación de texto. La automatización de las tareas de archivo y búsqueda de documentos liberan a los trabajadores de las tareas repetitivas que consumen mucho tiempo y a la vez reducen el monto de papeles en la oficina.

IV. UN AMBIENTE DE UN BUFETE DE ABOGACIA Y NOTARIADO.

En esta sección se describe el ambiente de trabajo en un bufete profesional de abogados y notarios, subrayando las tareas que se pueden automatizar utilizando bases de datos textuales.

A. Definición de conceptos de Informática Jurídica.

La Informática Jurídica, como su nombre lo indica, es la fusión de la ciencia de la Informática y el Derecho. La informática tiene un campo de acción muy extenso, tanto es así que día a día juega un papel importante en casi todas las actividades humanas. Por esta razón es que no es de asombrarse que en el área del Derecho exista un vínculo directo con la informática.

Existen diversas definiciones para el término de Informática Jurídica. Una de las definiciones más populares es la de Jordan Flórez (1985:49), la cual dice:

"Informática Jurídica es la utilización de los diferentes conceptos, categorías, métodos y técnicas propios de la Informática en el ámbito de lo jurídico. Se relaciona con la creación, flujo, clasificación, organización, sistematización y utilización de datos requeridos en la "producción" y/o en la administración de lo jurídico, así como en el estudio de las implicaciones o efectos que esta utilización produce en el Derecho y por consiguiente en la sociedad".

El objetivo de la Informática Jurídica consiste en racionalizar y organizar los datos jurídicos o los utilizados en el Derecho, de tal forma que sea posible su cualificación como Información; y que a su vez los profesionales o personas relacionadas con el Derecho puedan utilizar estos sistemas de una

manera fácil y comprensible.

Según su campo de acción, la Informática Jurídica se ha clasificado en tres grandes áreas: La Informática Jurídica Documental, La Informática Jurídica de Gestión y la Informática Jurídica Decisional.

1. La Informática Jurídica Documental.

Esta área se basa en el Documento jurídico, de ahí que se deriva su nombre. La estructura de estos sistemas está conformada por las técnicas de manejo de texto. Y su uso primordial se puede decir que es la consulta de los documentos por el usuario del Derecho.

Esta Informática Jurídica Documental se relaciona con la información legislativa, jurisprudencial y doctrinaria, así como también con bibliografía jurídica.

Dentro de la información legislativa encontramos los proyectos de ley, los proyectos de actos legales, la Constitución, las distintas leyes, los decretos, las ordenanzas, los acuerdos y las resoluciones.

La información jurisprudencial la conforman las sentencias de la Corte Suprema de Justicia, las sentencias del Consejo de Estado, las sentencias de los tribunales y las

de los Juzgados.

Por su parte, la información doctrinal cubre las doctrinas de tratadistas nacionales e internacionales.

La información bibliográfica jurídica cuenta con poderes, contratos, autos, escrituras, memoriales, etc.

2. La Informática Jurídica de Gestión.

Esta área de la informática jurídica trata, básicamente, con los procesos administrativos. El procesamiento de datos convencional abarca esta rama de la informática jurídica. Los sistemas que cubre esta área son: Sistemas de gestión de oficinas públicas o privadas; Sistematización del reparto judicial; Sistematización de notariado y registro; Sistematización del reparto penitenciario; Sistematización del registro de contravenciones; Sistematización de consultorios jurídicos.

3. La Informática Jurídica Decisional.

Esta área de la informática jurídica vincula la rama de la inteligencia artificial con la informática jurídica. Trata de los sistemas expertos para el Derecho. Un sistema experto se puede decir que simula por medio de la computadora el razonamiento y toma de decisiones de un experto en una rama específica, en este caso la rama del Derecho. Estos

sistemas se encuentran aún en desarrollo, por lo que no se profundiza en este documento.

En el área jurídica casi toda la información tiene un carácter netamente subjetivo. Esto quiere decir, que el problema básico de la informática jurídica no es la determinación de las fuentes de información ni su recolección, sino la dificultad de interpretación que cada persona pueda darle a la información.

B. Descripción del ambiente de trabajo en un bufete profesional de abogacía y notariado.

Así como el surgimiento de la Revolución Industrial cambió la agricultura por las fábricas a mitad del siglo XIX, la Revolución Electrónica promete cambiar la oficina actual de igual manera. Aceptar el reto de la Revolución Electrónica requiere de un aprendizaje de nuevas herramientas y nuevos sistemas para producir el trabajo legal, así como el conservar y mejorar varias de las habilidades del trabajo legal actual. El equipo de los abogados y notarios está cambiando de una simple pluma de escribir y pedazos de papel que cuestan centavos a aparatos sofisticados que cuestan miles de quetzales.

Hasta hace algunos años la oficina del abogado y notario no poseía el concepto empresarial que está empezando a tener en la actualidad. Las técnicas de mercadeo, publicidad, rentabilidad y gestión estaban totalmente excluidas del bufete jurídico.

El tremendo crecimiento del volumen de información a consultar, los costos de operación de una oficina de este tipo y el número de trabajos a desarrollar, ha reducido el tiempo disponible para realizar búsquedas de información con sistemas manuales. De aquí se deriva la necesidad de implantar técnicas de organización de negocios y la introducción de la mecanización y la informática para optimizar la operatividad de la oficina.

La automatización puede ser definida como la sistematización de la producción del trabajo por medio del mejoramiento de la actividad humana con tecnología moderna.

Las tareas que realiza un abogado diariamente son varias, pero en este trabajo mencionaremos únicamente las actividades que puedan de una u otra forma automatizarse.

Actualmente, en las oficinas de abogados en donde no existen computadoras el abogado y sus asistentes utilizan gran parte de su tiempo en la consulta de grandes volúmenes de información escrita para el desarrollo de sus tareas profesionales.

1. Control de Tiempo.

Las frases de Abraham Lincoln "El tiempo de un abogado es su patrimonio comercial" y "El asunto de las tarifas es importante ... Atendido apropiadamente, la justicia completa es aplicada a ambos el abogado y el cliente" Kubernan

(1980:s/n), nos muestran claramente la importancia que tiene el tiempo invertido en las actividades de un asunto propio de los abogados y notarios.

Entonces, podemos decir que el tiempo es el producto principal de venta que tienen los abogados, por lo que debería ser el mejor administrado en estas oficinas para una facturación justa a los clientes. Mientras mejor organizado esté el sistema de control de tiempo mejor será la productividad de la oficina del abogado y notario. Este aspecto es de los más relevantes e importantes para apreciar completamente el valor instrumental que la automatización de las tareas rutinarias tiene para el abogado y notario.

2. Facturación automatizada.

El control de tiempo junto con un sistema de registro de gastos acumulados para el trámite de los asuntos en estas oficinas forman parte de una de las aplicaciones principales en un bufete de abogados y notarios, la cual es la aplicación de facturación automática a los clientes.

La facturación apropiada a los clientes es de las tareas que requiere mucho cuidado y dedicación en el bufete de los abogados y notarios. Así como también es una actividad que se puede automatizar, fácilmente, con una mínima contribución por parte del abogado en la revisión de la cantidad que se le

cobra al cliente.

3. Sistema computarizado de inscripción, renovación y oposición de marcas, expresiones de propaganda y nombres comerciales.

Otra actividad que se realiza en un bufete de abogados y notarios es la inscripción, renovación y oposición de marcas, expresiones de propaganda y nombres comerciales en el Registro de la Propiedad Industrial. Esta actividad es perfectamente automatizable ya que los trámites que involucra son bastante repetitivos y salen, raramente, de lo normal. La automatización de esta actividad se puede realizar con un sistema de bases de datos convencional de formato fijo, aunque uno puede tener toda la información del bufete en una única base de datos de formato libre centralizada y accesible a todo el personal. Esto significa que una opción no excluye a la otra.

4. Protocolo del Notario.

Dentro de las tareas automatizables en un bufete de abogados y notarios podemos, igualmente, mencionar el Protocolo de los notarios, el cual tiene un formato bien definido y una función específica.

El protocolo es la reunión de todos los instrumentos públicos que un notario autoriza durante cada año, o sea, de todos los contratos que se formalizaron ante sus

servicios. Este protocolo tiene, a su vez, un índice. El índice de protocolo está regulado por los artículos 15 y 16 del código de notariado en vigor en Guatemala. Este índice es la ordenación por número, fecha, lugar, nombre de las partes que comparecieron en los contratos, tipo de contrato suscrito y el número de folio del protocolo del notario.

Las consultas que se pueden hacer en los protocolos pueden ser por ejemplo: cuáles son los instrumentos que se otorgaron en el año, o cuándo o cuál fue el documento que hizo alguna persona con otra en un momento determinado.

Existen muchos beneficios al automatizar el índice de protocolo, uno de ellos es si se quieren ver todos los instrumentos públicos que determinada persona ha otorgado ante los oficios de un notario determinado. Con ingresar el nombre se hace la búsqueda por año o por todos los datos que estén almacenados para ser consultados.

Otra utilidad en estos sistemas es la generación de información estadística, por ejemplo, determinar cuáles son todas las compra-ventas hechas en el año y pedir el dato numérico de cuántas encontró.

Otro beneficio es que hay muchos instrumentos públicos que los notarios realizan muy esporádicamente, por ejemplo un

fideicomiso testamentario o una escritura de medición de obligaciones sociales. En determinado momento se puede requerir esa escritura como modelo para hacer otra igual o parecida. Si se busca, manualmente, en el protocolo del notario se pierde mucho tiempo, mientras que si se tiene esta información en una base de datos textual, la consulta se puede realizar en un tiempo bastante corto sin mayor esfuerzo humano.

Una automatización en bases de datos textuales o de formato libre es la ideal para este tipo de aplicación por la naturaleza textual de los datos.

5. Sistema de Fechas de Vencimiento.

En un bufete de abogados y notarios, generalmente, los clientes confían en que se les lleve un seguimiento de las fechas en que venzan los contratos suscritos por ellos o las fechas en que se deban presentar casos a tribunales. Por esta razón debe existir un control estricto de estas fechas para que las acciones que se deban realizar se realicen con suficiente tiempo de anticipación y así cumplir con la eficiencia propia de una oficina jurídica.

Este tipo de automatización es muy frecuente en casi cualquier oficina jurídica en los Estados Unidos de Norte América, la cual se conoce como "Docket Control System".

Estos sistemas son prácticamente una agenda de fechas de vencimiento o de citas importantes, las cuales resultaría vergonzoso pasar por alto por un simple olvido.

Este sistema de control de fechas importantes puede ejecutarse en cualquier base de datos convencional de formato fijo, ya que la información manejada es casi uniforme y perfectamente codificable.

6. Contabilidad y sistemas de cuentas por cobrar y por pagar.

Toda oficina profesional debe, por ley, tener un sistema de contabilidad y también sistemas de cuentas por cobrar y por pagar para que funcione correctamente, así como también los sistemas de planillas y demás controles internos como inventarios, presupuestos, etc. Estos sistemas actualizan la contabilidad en general y permiten llevar de una forma detallada y precisa el control de ingresos y egresos. Un sistema de este tipo se enmarca perfectamente en una base de datos de formato fijo debido a la naturaleza fija del formato de los datos manejados.

Estas tareas son clásicas en la automatización de oficinas y empresas. Por esta razón únicamente mencionaremos que en un bufete de abogados debe existir este tipo de sistema computarizado.

7. Reproducción de documentos.

Esta actividad es automatizada por medio de los procesadores de palabras que existen actualmente. Los sistemas de procesamiento de palabras pueden convertir al abogado en una persona más productiva, particularmente en las oficinas en que se desarrollan sistemas y procedimientos de administración para la preparación de documentos.

Cuando un sistema de procesamiento de palabras se ha instalado y usado correctamente se reduce el tiempo de corrección y revisión de documentos, así como también la circulación de documentos se torna más rápida y la presentación se mejora. También se reducen los costos, ya que el personal de la oficina aumenta su productividad, pudiendo abarcar más asuntos que los que podría si utilizara sólo la máquina de escribir.

Existen varias características y funciones disponibles en los procesadores de palabras que incrementan significativamente la habilidad de producir documentos de alta calidad y libres de errores. Estas podríamos decir son las siguientes: la numeración automática de párrafos, la identificación de modificaciones dentro del documento, la generación automática de índices y tablas de contenido, la verificación de la ortografía dentro del documento, la existencia de índices cruzados para mantener conectada la

información por algún elemento específico, el control automático de unidad en los párrafos para prevenir que, por ejemplo, la última línea de un párrafo se imprima en otra hoja y la habilidad de fusionar un documento ya sea con información almacenada en otro documento o con información del operador de la pantalla.

8. Investigación Legal asistida por computadora.

Estos sistemas de recuperación de información para asistir el trabajo legal no deben cuestionarse como si fueran un sistema suficiente para recuperación de documentos legales, sino como si hacen una contribución significativa a todo el proceso disponible de la investigación. Estos sistemas forman parte de lo que definimos como Informática Jurídica Documental.

En el curso de su ejercicio profesional, cada abogado o notario va acumulando, constantemente, información en forma escrita de legislación, doctrinas y jurisprudencia. Esta información debe, entonces, consultarla el abogado, a veces por horas, hasta encontrar el o los documentos relevantes al caso que está tratando. El tiempo invertido en búsquedas de material textual se convierte en una tarea rápida y bastante precisa si se utiliza una base de datos textual. Lo único que debe determinar el usuario para hacer una búsqueda en la base de datos textual son los términos que se deberán buscar en la

información disponible.

El objetivo de estos sistemas es, entonces, responder a los requerimientos de información de los usuarios en cuestión de segundos, de tal forma que el usuario tenga a su disposición todo el material relevante a su búsqueda almacenado en la base de datos del computador.

Estos sistemas de soporte legal tienen una gran variedad de usos. A continuación se enumeran algunas de las tantas aplicaciones que se les puede dar: localizar casos con hechos semejantes al caso investigado; encontrar casos que cubran dos o más áreas de la ley; localizar casos sobre principios específicos; identificar las opiniones escritas de un juez particular; localizar casos que contengan alguna palabra o frase específica, o un grupo de palabras lógicamente relacionadas; localizar un caso en donde, únicamente, se conoce información limitada, como el nombre de la otra parte, etc.; localizar los casos más recientes en donde se mencione un artículo, tratado, decisión o estatuto; determinar cómo se ha aplicado la ley, en casos anteriores.

Estos sistemas también deben ayudar a prever las resoluciones judiciales que se puedan dar en un caso específico.

9. Comunicaciones.

Las comunicaciones son un elemento indispensable en una oficina de abogados y notarios. El medio de comunicación más popular hasta hace algunos años era el télex, conforme fue pasando el tiempo surgió con mucho éxito el faxsímil en donde se logran transmitir copias de los documentos o imágenes escritas por vía del teléfono. Actualmente, el teléfono es de los medios de comunicación más importantes en todo el mundo, y los adelantos tecnológicos en comunicaciones están encaminándose en ese sentido y en utilizar, también, líneas telefónicas dedicadas utilizando materiales más efectivos como la fibra óptica. En el mundo moderno existen ya varias computadoras grandes que tienen disponible grandes volúmenes de información que un usuario con una microcomputadora y un módem puede consultar por una tarifa razonable medida en función de los caracteres transmitidos y el tiempo de conexión. Los buzones electrónicos se han vuelto muy populares y permiten comunicaciones eficientes entre varias computadoras.

Lo importante de esta aplicación no es únicamente el hecho de facturar, efectivamente, al cliente las comunicaciones involucradas en el desarrollo del asunto tratado, sino también el ver la importancia que tiene ya en algunos países poder consultar grandes volúmenes de información legal almacenada y procesada en computadores

grandes. Este servicio de consulta de información está siendo la revolución de la segunda mitad de este siglo. Según Toffler (1984) en su libro "La tercera Ola", la información se convertiría en el capital de la revolución de este siglo.

C. Tareas automatizables utilizando Bases de Datos Textuales.

De las aplicaciones mencionadas en la sección B del capítulo IV "Descripción del ambiente de trabajo en un bufete profesional de abogacía y notariado", las que en este trabajo interesan son las que convenga ejecutar con una base de datos textual. En realidad, uno puede ejecutar todas sus aplicaciones en una base de datos textual para tener todos los papeles centralizados, así como también la información complementaria de una oficina de abogados y notarios. Este estado ideal de ejecución de todos los sistemas en una oficina, no puede ejecutarse perfectamente en la práctica debido a las limitaciones que aún siguen teniendo los paquetes de aplicaciones existentes en el mercado.

La parte práctica de este trabajo consta de una aplicación típica en una oficina de abogados y notarios. Dentro de las que mencionamos en el punto anterior, podemos contar con la automatización de índice de protocolos y la investigación legal asistida por computadoras.

Estas aplicaciones se podrían decir que son las que están más

al alcance de los bufetes guatemaltecos, ya que en otros países más avanzados casi todas las oficinas jurídicas cuentan con computadores grandes, o ya existe una base de datos nacional que puede ser consultada por cualquier usuario autorizado. En nuestro medio nos concentraremos en el ambiente para microcomputadores, el cual es mucho más accesible.

Con una microcomputadora se pueden ejecutar una gran variedad de aplicaciones utilizando una base de datos textual. En este trabajo nos concentraremos en la parte de la consulta a la legislación.

Prácticamente cualquier material que el abogado considere que es importante se puede ir almacenando en una base de datos de formato libre. La consulta de este material resulta risible comparado con un sistema manual de almacenamiento de documentos.

V. LA LEY DEL IMPUESTO SOBRE LA RENTA. UNA APLICACION ESPECIFICA

Esta parte del trabajo presenta el desarrollo y ejecución de la aplicación específica del manejo de leyes en el ambiente legal guatemalteco, utilizando para ello una base de datos textual o de formato libre. El desarrollo de esta aplicación es sobre la Ley del Impuesto sobre la Renta y abarca aspectos como la determinación de disposiciones existentes en esta Ley en relación a una palabra o a un grupo de palabras, la generación de índices para la ubicación precisa del texto que contenga cierta palabra y la combinación de un par de palabras en un mismo artículo de la Ley. Este sistema está orientado a un bufete de abogados de mediana escala y por consiguiente el sistema se desarrolló a manera de ejemplo para cualquier micro-computadora, con el popular sistema operativo DOS (Disk Operating System).

Se escogió un sistema de bases de datos textual que utiliza el modelo no jerárquico, explicado en la sección número III.D.2.b "Modelo No Jerárquico", debido a sus características favorables para una aplicación de este tipo. El tema escogido para desarrollar esta aplicación fue un tema de lo que es la informática jurídica documental, ya que trata, básicamente, de la consulta de la Ley del Impuesto sobre la Renta.

El sistema de bases de datos textual utilizado trabaja buscando en el texto completo que se ingrese al sistema. Podría haberse utilizado una base de datos creando resúmenes e índices de la información que se quiere consultar, pero esto implicaría

una interpretación individual de la Ley, además de un esfuerzo humano adicional. Según la experiencia y prácticas con profesionales del derecho las leyes deben consultarse en su forma original y no en una interpretación de la misma.

Un sistema de bases de datos textuales o de formato libre proporciona una magnífica solución al problema del manejo y localización de la información, tal como se explicó en las secciones anteriores. Una base de datos tradicional es ideal para manejar información que esté muy bien estructurada; una hoja electrónica es una buena solución para manejar datos numéricos; un procesador de palabras es bueno para crear información, pero deficiente para localizarla u organizarla; y, un sistema de base de datos textual puede procesar información que no esté estructurada, relacionada y planificada, siendo ideal para desarrollar aplicaciones como la que en este caso nos interesa.

Debido a la correspondencia imperfecta de las ideas con las palabras o frases del lenguaje utilizado en la mayoría de información textual, un sistema de este tipo adquiere también esta limitante. Las implicaciones que genera esta característica fueron explicadas con anterioridad en este trabajo.

A. Descripción de la Metodología a usarse para desarrollar la aplicación.

El sistema debe cubrir los siguientes objetivos:

- a) permitir el ingreso de información al sistema de una forma

- ordenada y consistente;
- b) permitir la actualización de la información ingresada ya sea por modificaciones de fondo o por correcciones a la información ingresada;
 - c) permitir el acceso a la información sin alterarla para cualquier tipo de consulta ya sea de una palabra específica o de un grupo de palabras relacionadas; y,
 - d) permitir la impresión de reportes de los textos encontrados que se relacionen con una consulta específica.

Para poder cubrir estos objetivos se diseñó el sistema en dos módulos principales:

MODULO I.

Ingreso y Mantenimiento de la información.

En este módulo el usuario podrá ingresar y corregir la información del sistema. Así como también el usuario del sistema podrá hacer las actualizaciones que se den para la información que ya esté ingresada en el sistema, dejando un registro de las alteraciones y cambios que sean realizados.

MODULO II.

Consultas y Reportes de la información del sistema.

En este módulo el usuario podrá hacer consultas en busca de la información que contenga una o varias palabras relacionadas con el tema o asunto que esté investigando. También podrá

imprimir reportes de la información que le interese.

El sistema se desarrolló utilizando la base de datos textual "Dayflo" y el procesador de macros "Smartkey". En el siguiente numeral se da una explicación del funcionamiento general del sistema Dayflo, con el objeto de mostrar lo poderoso que es un sistema de este tipo para el manejo moderno de cualquier información de formato libre y, en especial, para la aplicación desarrollada en este trabajo.

1. Diseño del paquete Dayflo:

El diseño del sistema DayFlo (base de datos de formato libre) está integrado por un sistema manejador de bases de datos y un procesador de texto. El Dayflo tiene un ambiente de operación semejante al de un escritorio que corresponde casi al mundo de papeles al cual estamos acostumbrados.

En DayFlo los registros están representados por pedazos de papel que pueden ser escritos, editados, tirados a la basura, amontonados, movidos alrededor del escritorio, o regresados al archivo. Como uno debe esperar, cualquier registro (posiblemente uno en blanco) que esté enfrente de uno puede ser modificado simplemente tecleando. De hecho, en DayFlo uno nunca deja el editor excepto, momentáneamente, para realizar alguna función utilitaria. Todo lo que hace uno es pasar registros diferentes a través del editor. Con

un simple teclazo uno se puede mover de registro en registro - editando o cortando y re-adequando, o hasta creando nuevos registros- sólo con teclear.

Este modelo de simulación de un ambiente de papeles (en contraste con uno orientado a formas) requiere un método nuevo de organización de bases de datos. Ya que un registro puede contener cualquier clase de información y puede no asemejarse a otro en la base de datos, el concepto de archivos diferentes dentro de la base de datos no tiene ningún sentido. Por esto, la base de datos no se puede describir como jerárquica o relacional. Piénsese en la base de datos de DayFlo como un único archivo en donde uno coloca todo junto y sin ningún orden preestablecido.

La recuperación de registros del sistema de base de datos es asociativa, basada únicamente en el contenido de cada registro. Por ejemplo, uno le puede decir al sistema, "recupere todo lo relacionado con planificación de negocios o finanzas que tenga que ver con Pedro García". El sistema entonces localiza una serie de registros, algunos de los cuales pueden ser memorándumes, cartas, reportes, o simplemente notas informales, y los coloca enfrente del usuario para que los pueda utilizar. El primer registro está listo para ser editado, mientras que los demás pueden ser accedados fácilmente al moverse por la pila que esté en

uso por medio de las teclas de movimiento de registros.

La arquitectura del DayFlo, como se muestra en el anexo "F", es simétrica y de terminación abierta. Esto significa que cualquier parte del sistema puede llamar a cualquier otra. Es más, todas las funciones del sistema regresan resultados uniformes como respuesta a estas funciones, lo cual permite que se pueda llevar un seguimiento de todas las operaciones ejecutadas. El sistema es en su esencia una matriz de tipos diferentes de operaciones que pueden realizarse sobre distintos tipos de datos y conformar una colección de módulos relativamente independientes.

El sistema incluye un grupo básico de funciones, tales como el editor, el organizador de escritorio, un ordenador, un procesador de encuentre-y-recupere, un manejador de pantallas y otros. También maneja tipos uniformes de datos, tales como texto, números y fechas, los cuales son entendidos en cualquier lugar del sistema. Nuevos tipos de datos y funciones también se pueden agregar.

El DayFlo es independiente del hardware, por lo que resulta fácil de transportar de máquina a máquina. La base está escrita en lenguaje assembly y el resto del programa está escrito en lenguaje C. Una de las características claves de la base, llamada Servicios de Ejecución, es una

memoria virtual especialmente diseñada y la capacidad de enlaces dinámicos. Este Servicio de Ejecución del DayFlo le permite hacer un uso óptimo de toda la memoria disponible por una rotación de módulos de programas conforme sean requeridos. Esto también permite llamar a cualquier función dentro del sistema sin la necesidad de reenlazar el sistema cada vez que se agregue una nueva.

El Procesador de Comandos y Respuestas maneja toda la comunicación con el usuario y actúa como un gran interruptor capaz de seleccionar los módulos del sistema apropiados para cada operación. Este Procesador realiza diálogos simultáneos con el usuario y el sistema para pedir y validar todos los parámetros necesarios para ejecutar cada operación. El control pasa a la rutina de servicio apropiada sólo cuando la validación se completa. Hasta este punto, se mantiene un seguimiento de la forma en que ha progresado el diálogo usuario-máquina. Este control central le permite a DayFlo no solo una ayuda sensitiva de contexto sino que también una manera rápida de recuperar errores.

Debido a que el diálogo se maneja de una forma centralizada, la percepción del usuario es la de un todo homogéneo, con muy poca estructura implicada.

La parte del programa que maneja la base de datos usa un

método de acceso sencillo, rápido y de llaves múltiples para el almacenamiento de todos los registros en un mismo archivo, el cual es el universo de datos de DayFlo. Este archivo es no jerárquico, y sus múltiples llaves permiten que los datos sean recuperados basados únicamente en su contenido. La construcción de índices opcional permite marcar aquellos elementos que sean de interés especial para optimizar su recuperación en las búsquedas. Los índices se almacenan en el mismo lugar de los registros de datos.

Se utiliza un diccionario para especificar los elementos que pertenecen al sistema, especificar su definición de tipo (usada para verificar el ingreso de datos, etc.) y determinar si están o no en el índice. El Manejador del Diccionario es responsable de proporcionarle información a otras partes del sistema sobre la definición de los elementos según se requiera. El diccionario reside normalmente en memoria. A la hora de limpiar el área de trabajo se almacena en un área fija al principio del archivo de la base de datos junto con otras tablas dinámicas de tal forma que sea posible reiniciar el trabajo en el punto donde se dejó el sistema al salir.

El Manejador del Área de Trabajo es responsable de determinar los registros que están en uso. También mantiene una tabla que muestra el contenido de todas las 20 pilas de documentos y el cesto de basura. Adicionalmente, provee las

funciones necesarias para mover registros a través del área de trabajo para su recuperación, orden o impresión.

DayFlo estructura sus registros de la base de datos de una manera única dentro de las bases de datos conocidas. Mientras que cada registro en DayFlo puede ser de cualquier forma o tamaño, internamente estos son almacenados como filas continuas de bytes con caracteres únicos de DayFlo. El conjunto de caracteres de DayFlo se diseñó especialmente para facilitar los tipos de operaciones realizadas en el sistema DayFlo. Los registros de DayFlo pueden entonces pasar entre los distintos componentes del sistema, que no requieren transformaciones en cada paso. Cada registro de DayFlo no solo contiene información de sus elementos sino que también su estructura y estado actual.

Si observamos el sistema Dayflo desde el punto de vista de una base de datos, un registro de este sistema es realmente una serie de elementos o campos, en donde cada elemento tiene un nombre y cualquier número de ocurrencias o valores. Cada registro tiene un identificador, el cual es siempre el primer elemento del mismo, siendo totalmente transparente para el usuario y el cual es utilizado para fines de recuperación de registros. Cada elemento o campo del registro también se marca con un carácter especial en hexadecimal. Existe también una marca especial, también en

hexadecimal, para indicar el fin del nombre del campo e indicar el inicio de un valor del campo. El cambio a otro valor del campo también se identifica con una marca especial. Como todos los registros en Dayflo no tienen una longitud específica, no existe límite en cuanto al número de campos y valores dentro del registro.

El resto de marcas especiales en los registros de Dayflo no se utilizan por el manejador de la base de datos, sino que por el editor de registros en las funciones de procesamiento de palabras. Existen marcas para los cambios de líneas, para los efectos especiales de video y todas las funciones del editor de textos.

El Manejador de Registros y el Editor están en constante comunicación para permitirle al usuario manejar la información en formato libre y a la vez actualizar la base de datos. El Manejador de la Base de Datos se comunica también con el Manejador de Registros para actualizar los índices de los datos y establecer las direcciones de los registros dentro de la base de datos.

El Dayflo, de por sí, es bastante complejo, y el espacio en memoria, relativamente, pequeño. Por esta razón el sistema se diseñó de tal forma que todos los dispositivos del computador entienden el formato interno del registro. No se

necesitan topes (buffers) adicionales ni se necesitan traducciones cada vez que se utiliza un registro, ya que cada carácter tiene su interpretación propia. Esto incrementa la eficiencia de ejecución del sistema.

El Manejador de Aplicaciones proporciona rutinas de aplicación general que permiten la incorporación de nuevas aplicaciones al sistema. El Manejador de Utilitarios provee una serie de rutinas interactivas para darle mantenimiento al Diccionario y la Base de Datos, incluyendo las operaciones de copias de respaldo de la base de datos y la comunicación de datos con sistemas externos.

Las acciones de todos estos manejadores son transparentes para el usuario del sistema. Lo que el usuario ve es una simulación de un ambiente de escritorio con veinte pilas de documentos o carpetas, los cuales se pueden usar para extraer información de la base de datos principal. El usuario se puede mover rápidamente de una pila a otra, de un registro a otro, extraer más información de la base de datos o regresar la información a la base de datos cuando la termine de utilizar.

El sistema Dayflo se diseñó para realizar una gran variedad de actividades de manipulación de texto. A pesar de que la base de la filosofía de DayFlo es la operación de

datos de formato libre, uno puede desarrollar fácilmente sistemas de mantenimiento de registros equivalentes a sistemas realizados con bases de datos de formato fijo.

B. Desarrollo de la aplicación "La Ley del Impuesto sobre la Renta".

El sistema, como se explicó en la sección V.A. "Descripción de la Metodología a usarse para desarrollar la aplicación de este trabajo", consta de dos módulos principales para cubrir los objetivos trazados.

La estructura o diseño del programa se esquematizó de la siguiente forma:

PRESENTACION DEL SISTEMA.

MENU PRINCIPAL DEL PROGRAMA.

- MANTENIMIENTO DE LA INFORMACION.
 - REVISAR, INGRESAR O MODIFICAR LOS ARTICULOS DE LA LEY.
 - ORDENAR LOS ARTICULOS DE LA LEY POR NUMERO.
 - REGRESO A MENU ANTERIOR.
- CONSULTAS Y REPORTES.
 - CONSULTA CON OPCION A IMPRESION DE LOS ARTICULOS DE LA LEY QUE CONTENGAN UNA O VARIAS PALABRAS O FRASES.
 - CONSULTA DE LOS ARTICULOS QUE CONTENGAN UN PAR DE PALABRAS EN EL MISMO ARTICULO DE LA LEY.
 - GENERACION DE UN INDICE DE LA LISTA DE PALABRAS QUE INGRESE EL USUARIO.

- REGRESO A MENU ANTERIOR.
- SALIDA DEL SISTEMA.

Dentro de cada una de las opciones se intercalaron pantallas de ayuda y mensajes para el usuario, haciendo el sistema bastante amigable.

Es importante notar que este sistema se presenta a manera de ejemplo, como una de las varias aplicaciones que se pueden desarrollar con un sistema de este tipo.

A continuación se presentan las diversas pantallas y vistas que tiene el usuario al utilizar el sistema, así como también una muestra de los reportes que permite imprimir el sistema.

Cuando el usuario quiere utilizar el sistema debe cerciorarse de que el sistema esté instalado, así como los programas y la base de datos, y que se encuentre en el directorio del sistema.

Para iniciar el sistema se debe teclear en la línea de comandos del sistema operativo: `ISR[Enter]`

Después de esto aparecerá en pantalla la presentación de la base de datos textual, se debe entonces presionar `[Enter]`, como la pantalla lo indica.

Luego se entra directamente a la presentación de esta aplicación. La pantalla que aparece es la que se muestra en la pantalla No.1. Para utilizar la aplicación el usuario deberá teclear lo que la pantalla respectiva le indique. En este caso [Alt]-U. En seguida aparece en el monitor la pantalla No.2, en la cual se debe presionar cualquier tecla para entrar a lo que sería el Menú Principal de la aplicación.

La pantalla No.3 es el Menú Principal de la aplicación. Allí el usuario puede escoger libremente la opción que quiera ejecutar con solo presionar las teclas que en esta se indican. De esta pantalla el usuario puede entrar al módulo I de mantenimiento de la información, al módulo II de consultas y reportes o salirse del sistema.

En todas las pantallas el usuario debe seguir las instrucciones que en ella se colocan para poder utilizar esta aplicación.

Si el usuario ingresa al módulo de mantenimiento de la información le aparece en pantalla la pantalla No. 4 que le permite editar los artículos de la Ley, ordenarlos o regresar al Menú Principal (pantalla No. 3).

En lo que es la edición de los artículos de la Ley, existe una pantalla de ayuda que el usuario puede llamar cuando quiera y

cuantas veces quiera. La pantalla No. 5 es la que se muestra al inicio de la edición de la información, y es la que indica cómo se puede llamar a la pantalla de ayuda. La pantalla No. 6 es la pantalla de ayuda para la edición de la información (ver pantalla No. 6). Cuando el usuario termina la edición de los artículos de la Ley puede regresar al menú anterior con las teclas [Alt]-E como lo indica la pantalla de ayuda. En esta parte de edición de la información se permite modificar los artículos ya existentes, borrar artículos o ingresar nuevos. Estas funciones son las que le permiten al usuario actualizar la Ley conforme las alteraciones o modificaciones formales que se le hagan a la misma, así como también revisar y corregir la información ya ingresada.

La segunda opción de la pantalla No.4, o sea ordenar los artículos de la Ley, sirve para ordenar los artículos de la Ley conforme la numeración de artículos, incisos y sub-incisos. La pantalla que aparece en la pantalla cuando se usa esta opción es la pantalla No. 7. Este proceso de ordenamiento es ideal realizarlo antes de editar la información, para poder ubicar con más facilidad los artículos de la Ley que se necesiten editar, ya que las consultas y reportes cambian el orden de los artículos de la Ley.

La tercera opción regresa el control del sistema al menú principal (pantalla No. 3).

La segunda opción del Menú Principal es el módulo de consultas y reportes. Cuando el usuario entra en esta parte se muestra en el monitor el Menú de Consultas y Reportes (pantalla No. 8). En este menú se permiten realizar tres tipos diferentes de consultas.

La primera es una consulta de los artículos de la Ley que contengan una o varias palabras o frases en el texto de los mismos. En esta consulta son recuperados todos los artículos que contengan por lo menos una de las palabras o frases que el usuario indique. La pantalla No. 9 es la que se muestra al inicio de esta consulta, la cual indica la forma de llamar la pantalla de ayuda para ingresar las palabras o frases a buscar. [Alt]-C es la combinación de teclas que se deben presionar para ver la pantalla de ayuda. Esta pantalla de ayuda es la que se muestra en la pantalla No. 10. El máximo de palabras o frases que puede ingresar el usuario son diez. Al ingresar la última el usuario debe presionar las teclas [Alt]-S para que se realice la búsqueda de la información.

Cuando el sistema encuentra la información requerida por el usuario la coloca en un espacio que le permite el acceso al usuario. La pantalla No. 11 es la que aparece en el monitor. En esta pantalla se muestra un menú de opciones para ver la información en pantalla, imprimirla en un reporte o abandonar esta consulta y regresar el menú de consultas y reportes.

Si el usuario escoge ver la información en pantalla el sistema le presentará una pantalla de ayuda para poder revisar los artículos encontrados (ver pantalla No. 12). Al terminar la revisión en pantalla de la información el usuario debe presionar la teclas [Alt]-N para regresar al Menú de Resultado de la Consulta (pantalla No. 11).

Si se escoge la opción de imprimir la información encontrada aparecerá en pantalla el mensaje de alistar la impresora para poder imprimir (ver pantalla No. 13). Cuando la impresora esté lista con papel tamaño carta se debe presionar cualquier tecla y el sistema procederá a imprimir el reporte respectivo (ver Reporte No. 1). Al terminar la impresión el control del sistema regresa al Menú de respuesta de Consulta (pantalla No. 11). La tercera opción en este menú es para regresar al Menú de consultas y reportes, pantalla No. 8.

La segunda consulta definida en el sistema es la de encontrar los artículos que contengan dos palabras o frases en el texto del mismo artículo. Con la pantalla No. 14 el sistema le pide al usuario ingresar la primera expresión a buscar y luego con la pantalla No. 15 la segunda expresión. Se puede usar la pantalla de ayuda [Alt]-C (pantalla No. 10) para utilizar expresiones más complejas. Después de ingresar las dos expresiones el sistema se coloca en la pantalla de Resultado de Consulta (pantalla No. 11), y se procede igual que en la consulta

No. 1. El reporte que se imprime es el reporte No. 2. Este reporte es igual al No. 1, pero la diferencia está en el proceso de la generación del mismo.

La tercera opción del Menú de Consultas y Reportes es el reporte de la generación de un índice de los artículos de la Ley en base a la lista de palabras que el usuario ingrese al sistema. La pantalla No. 16 aparece en la pantalla al ingresar a este reporte. En ella se indica que con las tecla [Alt]-O se pide ingresar otra palabra y con [Alt]-T se termina el ingreso de palabras y el sistema empieza a generar el índice respectivo. Nótese que las palabras no tienen que ingresarse en orden alfabético ya que el sistema se encarga de ordenarlas a la hora de imprimir el reporte. Luego se muestra en el monitor la pantalla No. 17, la cual le da al usuario las instrucciones para poder concluir la generación del índice. El número de palabras de la lista está siempre en la última línea de la pantalla. La pantalla No. 18 debe aparecer en la pantalla si las instrucciones se ingresaron correctamente.

Cuando el sistema termina la generación del índice coloca en la pantalla el mensaje de alistar la impresora, y nuevamente el usuario deberá colocar papel tamaño carta y conectar la impresora para poder proceder a imprimir el reporte No. 3.

La última opción del Menú de Consultas y Reportes regresa el

control del sistema al menú principal y la última opción del menú principal es la salida del sistema. Cuando el usuario sale del sistema se muestra en el monitor la pantalla No. 19.

A continuación se presentan las pantallas mencionadas en esta explicación del funcionamiento del sistema, seguidas de los reportes también mencionados.

Pantalla No. 1

BIENVENIDOS AL SISTEMA

LEY DEL IMPUESTO SOBRE LA RENTA

(Para empezar presionar la tecla [Alt]-U.)

DayFlo: Press F1 for help

Pantalla No. 2

S I S T E M A D E C O N S U L T A

LEY DEL IMPUESTO SOBRE LA RENTA

Presione cualquier tecla para continuar

DayFlo: Press F1 for help

Menu Principal

SISTEMA DE CONSULTAS
LEY DEL IMPUESTO SOBRE LA RENTA

- [F1] Mantenimiento de la Informacion
- [F2] Consultas y Reportes
- [F3] Salida del Sistema

DayFlo: Press F1 for help

Pantalla No. 4

Menu de Mantenimiento

SISTEMA DE CONSULTAS
LEY DEL IMPUESTO SOBRE LA RENTA

- [F1] Revisar/Modificar Articulos
- [F2] Ordenar la Ley
- [F3] Regreso a Menu anterior

ayFlo: Press F1 for help

Pantalla No. 5

Ley IMPUESTO SOBRE LA RENTA

Capitulo I.- OBJETO Y CAMPO DE APLICACION

Articulo 1	Epigrafe	Objeto
Inciso	Sub-Inciso	

Texto El impuesto sobre la Renta es anual y recae sobre las rentas o utilidades que provengan de la inversión de capital, del trabajo o de la combinación de ambos y sobre los incrementos de patrimonio neto, que se produzcan en el período anual de imposición, generado por rentas gravadas no declaradas, más, en su caso, los gastos establecidos en esta ley y su reglamento.

Edicion Articulos

DayFlo: Press F1 for help

AYUDA ... [Alt]-A

Enter para seguir.

Pantalla No. 6

AYUDA EDITOR

[Alt]-A AYUDA

[F9] Regresar Registro

[F10] Avanzar Registro

[F6] Avanzar Campo

[Shift]-[F6] Regresar Campo

[Alt]-I Agregar Registro

[Alt]-B Borrar Registro

[Ctrl]-[End] Parrafo sig.

[Ctrl]-[Home] Parrafo ant.

[Alt]-E SALIDA

ayFlo: Press F1 for help

Pantalla No. 7

Proceso de Ordenamiento

El sistema procederá a ordenar la
Ley por No. de Artículo e Inicisos.

Favor esperar hasta regresar al MENU.

G R A C I A S

<pulse cualquier tecla>

Pantalla No. 8

Menu de Consultas y Reportes

SISTEMA DE CONSULTAS
LEY DEL IMPUESTO SOBRE LA RENTA

- [F1] Articulos que contengan 1 o varias palabras o frases (Max.10).
- [F2] Articulos que contengan 2 palabras o frases en el mismo.
- [F3] Generar Indice de una lista de palabras.
- [F4] Regreso a Menu anterior

ayFlo: Press F1 for help

Pantalla No. 9

~~Frases a buscar~~
Ingrese las frases
que quiere buscar.
Alt-C AYUDA

Texto contains what? _____
Value required _____
Help Status _____
WILD _____

Pantalla No. 10

AYUDA INGRESO DE FRASES A BUSCAR

Alt-C AYUDA

Ingrese las frases o palabras separadas por Enter. Los siguientes simbolos pueden usarse para expresiones complejas:

- * representa un cadena de cualquier caracter excepto espacio y de longitud variable.
- @ representa cualquier letra sola.
- # representa cualquier digito solo.
- representa cualquier caracter excepto espacio.
- ? indica que el caracter anterior es opcional.

Maximo de frases = 10

Al ingresar la ultima: [Alt]-S

Pantalla No. 11

Resultado de la Consulta

- [F1] Ver Articulos en Pantalla
- [F2] Imprimir Articulos encontrados
- [F3] Regresar a Menu Anterior

Pantalla No. 12

REVISION EN PANTALLA

[F10] Avanzar Registro

[F9] Regresar Registro

↑↓ Línea arriba y abajo respectivamente

←→ Caracter a la izquierda y derecha respectivamente

[Alt]-N CUANDO TERMINE

Pantalla No. 13

ALISTE LA IMPRESORA

Presione cualquier tecla cuando
la impresora este lista.

Pantalla No. 14

INGRESE LA PRIMERA FRASE

Ingrese la primera frase a buscar.

<Presione cualquier tecla para seguir>

Pantalla No. 15

~~INGRESE LA SEGUNDA FRASE~~

Ingrese la segunda frase a buscar.
<Presione cualquier tecla para seguir>

Pantalla No. 16

PALABRA

INGRESO LISTA DE PALABRAS

[Alt]-O Para ingresar otra palabra

[Alt]-T Cuando Termine

Pantalla No. 17

PRESIONE LAS SIGUIENTES TECLAS:

[Ctrl]-[F5]

C

A

R

No. Palabras de la lista

[Enter]

[Ctrl]-[F5]

C

E

[F5]-r

[Ctrl]-[F5]

C

E

[F5]-t

[Ctrl]-[F5]

[Enter] [Enter] [Alt]-Z

Pantalla No. 18

Defining [Alt Z] Enter definition ([^F5] for Definition Menu)

<R3><E[SS r]><E[SS t]>

Pantalla No. 19

GRACIAS POR HABER UTILIZADO EL SISTEMA

Utilice el Sistema cada vez que quiera consultar la Ley del Impuesto sobre la Renta.

HASTA LA VISTA.

Reportes generados por esta aplicación de la Ley
del Impuesto sobre la Renta.

Reporte No. 1

ARTICULOS LEY IMPUESTO SOBRE LA RENTA

ARTICULOS ENCONTRADOS EN LA CONSULTA REALIZADA

III.- RENTA DE FUENTE GUATEMALTECA

ARTICULO: 4 b)
Principios Generales (Art.1 / Decreto 95-87)

Los producidos por capitales, bienes o derechos invertidos, o utilizaados en el país, comprendiéndose dentro de los mismos los intereses de depósitos o de préstamos de dinero, de títulos, de bonos, de notas y de otros valores emitidos por personas de derecho público o privadas domiciliados en el país, utilidades o participaciones sociales, distribuidas o no, por personas jurídicas domiciliadas en el país; arrendamientos de bienes muebles; regalías; y otras rentas que revisten características similares;

IV.- DE LAS EXENCIONES

ARTICULO: 6 ñ)
Rentas exentas. (Art.2 / Decreto 95-87)

Los intereses producidos por depósitos constituidos en instituciones bancarias y financieras que operan en el país; así como los intereses producidos por bonos y títulos de crédito de reconocida solidez, emitidos o garantizados por el Estado, las entidades públicas, las instituciones financieras, los bancos que operan en el país y las empresas privadas cuyas emisiones califique como de primer orden la Comisión de Valores a que se refiere el artículo 112 de la Ley Orgánica del Banco de Guatemala, incluyendo las rentas o beneficios obtenidos en fideicomisos constituidos por valores garantizados de pólizas de seguros de vida.

IV.- DE LAS EXENCIONES

ARTICULO: 6 o)
Rentas exentas.

Intereses sobre bonos, notas o letras de tesorería emitidos por el Estado, sus instituciones autónomas, descentralizadas, y por las Municipalidades;

XI.- DE LA RENTA NETA

ARTICULO: 39 c)
Costos y Gastos Deducibles.

ARTICULOS LEY IMPUESTO SOBRE LA RENTA

Los gastos necesarios para las explotaciones agropecuarias, tales como: los efectuados por concepto de roturación y preparación de tierras, siembras, resiembras, riego, forrajes, plantas, semillas, abonos, fertilizantes, productos sanitarios, pastaje, conservación de cercas, podas, alimentación y crianza del ganado utilizado en la agricultura y demás gastos semejantes, en la proporción que corresponda a las cosechas que se obtengan en cada ejercicio económico contable; en los casos de comercialización de ganado, constituye costo deducible, la compra y gastos de engorde del mismo;

Reporte No. 2

ARTICULOS LEY IMPUESTO SOBRE LA RENTA

ARTICULOS ENCONTRADOS EN LA CONSULTA REALIZADA

III.- RENTA DE FUENTE GUATEMALTECA

ARTICULO: 4 b)
Principios Generales (Art.1 / Decreto 95-87)

Los producidos por capitales, bienes o derechos invertidos, o utilizaados en el país, comprendiéndose dentro de los mismos los intereses de depósitos o de préstamos de dinero, de títulos, de bonos, de notas y de otros valores emitidos por personas de derecho público o privadas domiciliados en el país, utilidades o participaciones sociales, distribuidas o no, por personas jurídicas domiciliadas en el país; arrendamientos de bienes muebles; regalías; y otras rentas que revisten características similares;

Reporte No. 3

INDICE DE LA LEY DEL IMPUESTO SOBRE LA RENTA

Indice Generado de la Lista de Palabras Ingresada

ahorro	Artículo	INCISO	SUB-INCISO
	39	r)	
arrendamiento	Artículo	INCISO	SUB-INCISO
	4	a)	
	4	b)	
	5	i)	
	16		
	69		
	81	b)	
bonos	Artículo	INCISO	SUB-INCISO
	4	b)	
	6	ñ)	
	6	o)	
	39	c)	
capital	Artículo	INCISO	SUB-INCISO
	1		
	4	b)	
	5	c)	
	6	a)	
	20		
	21		
	26		
	27		
	29		
	30		
	31		
	39	r)	
	41	d)	
	41	f)	
	51		
	81	b)	
	82		
	83	f)	
	111		
	117		
	118	b)	

INDICE DE LA LEY DEL IMPUESTO SOBRE LA RENTA

interés

Artículo	INCISO	SUB-INCISO
4	b)	
5	c)	
6	g)	
6	n)	
6	ñ)	
6	o)	
6	p)	
8		
32		
39	11)	
39	m)	
40		
41	d)	
48	b)	
64		
65		
66		
67		
69		
76		
77		
78		
80	c)	
80	e)	
80	m)	
80	n)	
81	b)	
85		
97		
99		
100		
107		
108		
110		
118	b)	

Como se dijo anteriormente, el sistema se desarrolló con la base de datos textual DayFlo y el procesador de macros (grupos de instrucciones) Smartkey. Este procesador de macros es un programa residente en memoria cuya herramienta principal es la programación del teclado. También se utilizaron los comandos de "Log" del sistema de base de datos textual para complementar los programas. Los programas de esta aplicación se listan en el Anexo "A".

Para poder realizar esta aplicación hubo que trabajar en la base de datos textual definiendo la forma de ingresar la información de la Ley del Impuesto sobre la Renta en el sistema. Se definieron los elementos necesarios en el Diccionario del DayFlo así como los reportes y pantallas de ingreso de la información.

Lo que se tomó en cuenta para ingresar la Ley fué la forma en que está estructurada la misma. La Ley del Impuesto sobre la Renta tiene un esquema ya definido el cual podemos decir que es el siguiente:

- Capítulos
 - Artículos
 - Incisos
 - Sub-Incisos

En base a este esquema y a la información que contiene esta

Ley, se definieron en el diccionario del sistema los siguientes elementos: **Ley IMPUESTO SOBRE LA RENTA, Capítulo, Artículo, Epígrafe, Inciso, Sub-Inciso y Texto.** El Anexo "B" muestra los elementos del diccionario.

El elemento **Ley IMPUESTO SOBRE DE RENTA** sirvió para diferenciar la información de esta Ley con la del resto del sistema. El **Epígrafe** es el texto escrito en la numeración de los artículos. Este separó de lo que es en sí el número de artículo para poder ordenar posteriormente la información por el número de artículo. El elemento llamado **Texto** es el contenido de cada uno de los artículos de la Ley. En este es en el cual se realizan la mayoría de las búsquedas de información. Los elementos **Artículo, Inciso y Sub-Inciso** sirven para ordenar la Ley.

Se dispuso colocar toda la información en cada pantalla de ingreso para fines de referencia inmediata del texto encontrado. El Anexo "C" muestra la pantalla de ingreso definida para este sistema.

VI. EVALUACION DEL SISTEMA DESARROLLADO EN ESTE TRABAJO Y OTRAS APLICACIONES UTILIZANDO BASES DE DATOS TEXTUALES.

En este capítulo se presenta una evaluación de la utilidad que pueden tener las aplicaciones que se ejecuten usando bases de datos textuales en el campo legal guatemalteco. Esta evaluación consta de dos partes principales: la primera versa sobre las ventajas y desventajas encontradas en la aplicación desarrollada en este trabajo, relativas al sistema de base de datos textual DayFlo; la segunda parte es una evaluación más general sobre la utilidad que pueden tener los sistemas de bases de datos textuales, como la aplicación del Impuesto sobre la Renta, en un bufete de abogados guatemalteco.

En los sistemas de búsqueda de documentos en computadora, como en el desarrollado en este trabajo existen una serie de factores que influyen para el éxito o fracaso del mismo.

A. Ventajas y beneficios de la utilización del sistema Dayflo.

Conforme la experiencia práctica adquirida al desarrollar el sistema de La Ley del Impuesto sobre la Renta utilizando la base de datos textual Dayflo, se pueden obtener las siguientes conclusiones en cuanto a ventajas y beneficios del uso de este sistema:

- a) Uno de los beneficios más evidentes es el ahorro de tiempo que se obtiene al utilizar una base de datos textual con la información que se necesita encontrar, en lugar de que la búsqueda se realice manualmente por un individuo, corriéndose incluso el riesgo de cometer errores humanos en la búsqueda,

con lo que se deteriora aún más el proceso de encontrar la información requerida.

- b) Otra ventaja que se tiene es el ahorro de espacio físico para almacenar la información que se quiera tener disponible para ser consultada en cualquier momento.
- c) El formato de los registros de la base de datos es libre, o sea que un registro no tiene porqué ser igual a cualquier otro. La longitud de los campos en los registros es variable y en cualquier momento se puede agregar cualquier campo a cualquier registro sin tener que reformatear toda la base de datos.
- d) El proceso de introducir la información es muy eficiente, ya que tiene incorporadas las funciones de un procesador de palabras.
- e) Permite tener varios valores por cada campo y colocar la información en cualquier lugar.
- f) Uno de los beneficios más importantes es la capacidad de recuperar registros en base al contenido de los mismos y no en base a criterios predefinidos.

B. Desventajas y limitaciones de la utilización del sistema Dayflo.

En cuanto a las limitaciones y desventajas encontradas al desarrollar la aplicación en el sistema DayFlo, se puede decir que existen varias que son importantes mencionar para que sirvan de referencia y guía para cualquier aplicación que se quiera realizar en un futuro ya sea con Dayflo o con cualquier otro

sistema de base de datos textual.

Entre las limitantes podemos mencionar que el sistema tiene un límite máximo de 5,000 documentos que pueden ser ingresados a la base de datos. Este límite máximo hay que tomarlo en cuenta para determinar si es o no aceptable de acuerdo a la cantidad de documentos que se quieran almacenar en la base de datos de la aplicación que se va a realizar.

Otra limitante tanto del sistema Dayflo como del procesador de macros Smartkey, que se usaron para desarrollar la aplicación de este trabajo, es que carecen de mecanismos de control de situaciones condicionales en el funcionamiento de los mismos cuando se quiere programar alguna aplicación que lo requiera. Por ejemplo a la hora de querer realizar búsquedas un número variable de veces dependiendo de los requerimientos del usuario no existe un mecanismo de control propio en estos dos sistemas. Para poder programar el reporte de índice de la lista de palabras ingresada por el usuario, en la aplicación de la Ley del Impuesto sobre la Renta, hubo que usar algunas técnicas de programación más complicadas y no tan amigables para el usuario, ya que el usuario debe indicar dentro de una definición del procesador de macros el número total de palabras ingresadas para generar el índice. Esto es inconveniente, a pesar de que el total de palabras aparece en la parte inferior derecha de la pantalla en el momento de tener que ingresar este número. Se

usaron también archivos de texto ASCII para poder colocar el resultado de las diferentes búsquedas en un sólo reporte general.

El tiempo de respuesta es otro factor importante que hay que tomar en cuenta para establecer cuál es la computadora más apropiada para ejecutar la aplicación desarrollada en Dayflo. El costo de la computadora y la velocidad de proceso de la misma van vinculados, y dependiendo de la aplicación y del tiempo de respuesta requerido así se debe tomar la decisión.

Otro aspecto que vale la pena mencionar no como una desventaja o ventaja, es el hecho de escoger el sistema de base de datos textual que permita los métodos de búsqueda necesarios a la aplicación que se quiera ejecutar y no sólo el de búsqueda en el texto completo como la que tiene el Dayflo. Para la consulta de leyes, sin embargo, resulta muy conveniente el realizar las búsquedas en el texto completo por el hecho de consultar las leyes en su forma pura y no interpretada. Para otras aplicaciones puede ser muy importante poder tener resúmenes de la información a consultar y también enlaces fijos pre-establecidos entre la información guardada en la base de datos.

Existen otros sistemas de bases de datos textuales que tienen la capacidad de presentar el resultado de las búsquedas en base al "peso" que tenga(n) la(s) palabra(s) buscada(s) en cada uno de los documentos recuperados. El sistema Dayflo carece de

esta capacidad.

Para escoger el sistema de base de datos textual para desarrollar la aplicación de este trabajo se recabó información de varias bases de datos textuales que existen actualmente para desarrollar estos sistemas, y se encontraron varias con características mejores y peores que el Dayflo. Se desarrolló la aplicación en Dayflo por la disponibilidad de la misma y por su aplicación en ambientes de micro-computadoras, por lo que resulta más accesible para el medio guatemalteco.

Otra limitante que tiene el sistema Dayflo es que es un sistema monousuario que no corre en redes locales de microcomputadoras ni en ambientes multiusuarios. Cuando existen varios individuos queriendo realizar búsquedas de información en el sistema al mismo tiempo, se crea un conflicto ya que sólo uno a la vez puede utilizar el sistema. Lo ideal es que un sistema de este tipo pueda ser utilizado al mismo tiempo por varios usuarios si la demanda de información es muy alta.

C. Evaluación de la utilidad del sistema DayFlo en una firma de abogados y notarios.

Esta sección presenta un análisis de la utilización de las bases de datos textuales en un bufete de abogados y notarios guatemalteco. Esta sección consta de tres partes principales. En la primera se hace una descripción y definición del ambiente y

situaciones analizadas para poder desarrollar el análisis de utilidad de estos sistemas de bases de datos textuales en un bufete de abogados y notarios. La segunda parte describe al detalle el análisis realizado así como también se presentan los resultados obtenidos. En la última parte se presentan las conclusiones derivadas de los resultados obtenidos en la investigación o análisis de utilidad que se llevó a cabo.

1. Descripción y definición del material a analizar.

En esta parte se hace la exposición del ambiente analizado para realizar la evaluación de utilidad de estos sistemas.

El análisis de utilidad de estos sistemas de bases de datos textuales se desarrolló en un bufete de abogados y notarios guatemalteco. Este bufete tiene funcionando varias aplicaciones utilizando bases de datos textuales y está conformado por siete abogados principales, personal legal asistente y demás personal de apoyo.

Para el análisis de este trabajo se escogieron 3 bases de datos de información diferente para poder evaluar la frecuencia de uso y el resultado de las búsquedas realizadas en estos sistemas.

El tamaño en número de registros de estas bases de datos

textuales y su clasificación se listan a continuación:

Tabla 6.1

Información tomada en cuenta en la evaluación

BASES DE DATOS TEXTUALES	No. REGISTROS
1.- Leyes:	
- Ley del Impuesto sobre la Renta	294
- Ley del Impuesto al Valor Agregado	147
- Ley del papel Sellado y Timbres Fiscales	149
- Constitución Política	421
2.- Indice de las Gacetas de los Tribunales (Procesos y Doctrinas)	6,130
3.- Indice de Protocolo de los Notarios del bufete	4,392

Estas bases de datos están organizadas de la siguiente manera: Las leyes se encuentran ingresadas todas en una misma base de datos, aunque en lugares y formas de registro diferentes para poder utilizarlas en conjunto o independientemente. Las Doctrinas de las Gacetas de los Tribunales se encuentran agrupadas en tres grupos que guardan, cada uno, veinte años de esta información, o sea, el primer grupo tiene la información de 1930 a 1949, el segundo de 1950 a 1969 y el tercero de 1970 a 1989. Esto está así por la limitación de número máximo de registros en cada base de datos. El índice de protocolos está ubicado

todo en una misma base de datos, en donde se diferencian los protocolos de cada notario y de cada año.

Estas aplicaciones se ejecutan utilizando la base de datos textual Dayflo. Dayflo es una base de datos textual típica que nos sirve como prototipo para poder visualizar el comportamiento de un sistemas de búsqueda de texto completo en el ambiente en cuestión. Es un sistema que nos proporciona las herramientas necesarias para realizar búsquedas en los textos en donde una o varias palabras específicas aparezcan ya sea en su forma pura o utilizando combinaciones booleanas.

Lo que se pretende evaluar en este análisis es la frecuencia de utilización que puede tener una base de datos textual en un bufete de abogados y notarios en Guatemala, así como también medir la precisión de los resultados obtenidos por las búsquedas realizadas.

Para medir la frecuencia de utilización de estos sistemas, se tomaron tres meses como el período de muestra, para determinar cuántas veces se realizaron consultas a estas 3 bases de datos. Las consultas a las bases de datos textuales de esta evaluación las formularon ya sea los abogados o el personal auxiliar del abogado. Por cada una de las consultas se llenó el formulario llamado SOLICITUD DE

INFORMACION, el cual se muestra en el Anexo "E".

En este formulario se apuntó a la hora de solicitar una consulta la siguiente información: la fecha de solicitud, el nombre de la persona solicitante, el código del expediente, el nombre de la base de datos donde se quería buscar, y las palabras o frases que se querían buscar indicando también si estas palabras debían estar todas en el mismo registro al mismo tiempo, o si con una que estuviera bastaba para recuperar el registro.

Se realizaron 89 consultas en los 3 meses. Los abogados o sus asistentes formulaban la pregunta y luego el personal calificado para utilizar el sistema DayFlo la convertía a una forma adecuada para que el sistema obtuviera la información buscada. Las búsquedas se realizaron en la base de datos hasta obtener un resultado satisfactorio de documentos encontrados, luego se le entregaban al abogado o asistente interesado y éste determinaba si se necesitaba o no un refinamiento a la búsqueda realizada. Si se necesitaba un refinamiento, se procedía a realizar de nuevo la o las búsquedas en el sistema hasta lograr recuperar los documentos a satisfacción del interesado. Sólo hasta que la persona que formuló la pregunta original estuviera satisfecha de los documentos recuperados, era que se establecía la consulta como concluida, y se procedía a

llenar el resto del formulario anotando el número de registros encontrados, la fecha de entrega, el total de documentos relevantes recuperados, y si se encontró o no información adicional a la prevista. Luego se calculó la precisión y se determinó si se había logrado encontrar información adicional a la prevista.

Las medidas para evaluar la efectividad de un sistema de búsqueda de documentos son precisión, recolección y desacierto, tal como se mencionó en la parte teórica de este trabajo. En esta evaluación únicamente logramos medir la precisión debido a que las variables involucradas en su cálculo se pudieron determinar fácilmente en todas las búsquedas, mientras que para poder medir la recolección y el desacierto se hubiera tenido que conocer en cada búsqueda realizada el número total de documentos relevantes en la base de datos consultada y el número total de documentos irrelevantes respectivamente. Para calcular la precisión únicamente había que conocer el total de documentos recuperados y el total de esos documentos que fueran relevantes a la pregunta realizada.

Otro parámetro importante para evaluar estos sistemas de bases de datos textuales fué el haber establecido en cada una de las búsquedas si se había encontrado información adicional no prevista antes de formular la pregunta de la

búsqueda.

La precisión se calculó generando el porcentaje de documentos relevantes recuperados del total de documentos recuperados. En los casos en que se realizaron varias búsquedas al sistema sobre una misma pregunta, se tomó la unión de los documentos encontrados en cada búsqueda como el total de documentos recuperados, o sea, si un mismo documento aparecía en varios resultados de las búsquedas realizadas, sólo se tomaba una vez en el total de documentos recuperados.

2. Desarrollo del análisis de utilidad del sistema de bases de datos textual Dayflo en el ambiente de un bufete de abogados y notarios.

En la sección anterior se dió una presentación de lo que se hizo para poder evaluar la utilidad de estos sistemas de bases de datos textuales en un bufete de abogados y notarios en Guatemala.

El total de consultas tomadas en cuenta para este estudio fueron 89 en un lapso de 3 meses de operación. Conforme los datos recabados en el formulario diseñado para utilizar estos sistemas se obtuvo la información necesaria para evaluar la frecuencia de uso, la precisión en las búsquedas y si se encontró o no información adicional a la prevista. En el Anexo "D" se muestra el resultado obtenido

de las 89 consultas realizadas ordenadas por base de datos consultada.

a) Frecuencia de utilización.

Para medir la frecuencia de utilización de estos sistemas se tomaron como base los tres meses de consultas mostradas en los resultados de este estudio. Luego se contaron las veces de utilización de estas bases de datos en este lapso de tiempo generando el siguiente cuadro de resultados:

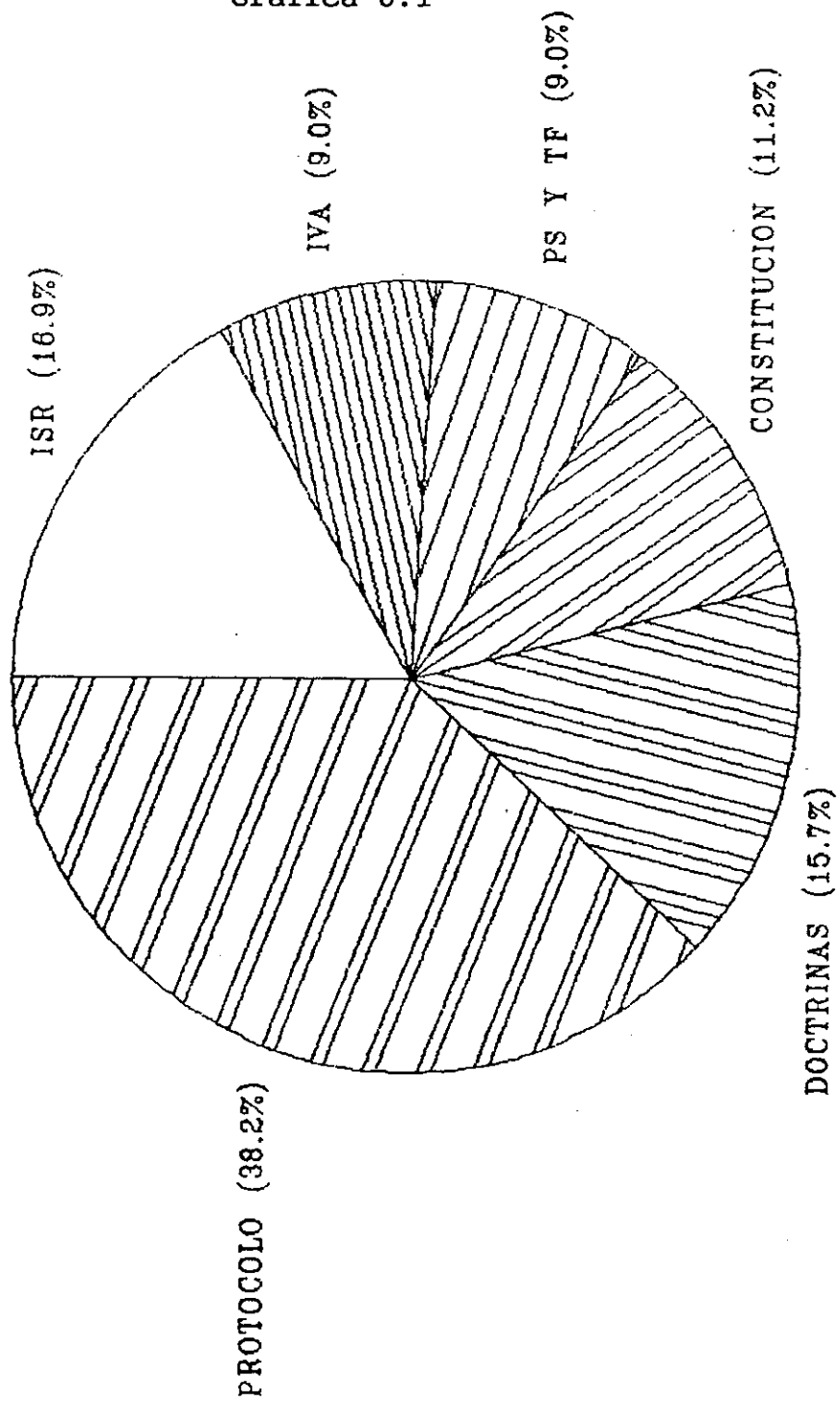
Tabla 6.2
Frecuencia de utilización

Base de Datos Consultada	Número de Consultas	Porcentaje
LEY IMPUESTO SOBRE LA RENTA	15	16.854%
LEY IMPUESTO AL VALOR AGREGADO	8	8.989%
CONSTITUCION POLITICA	10	11.236%
LEY PAPEL SELLADO Y TIMBRES	8	8.989%
GACETAS DE LOS TRIBUNALES	14	15.730%
PROTOCOLOS	34	38.202%
T O T A L	89	100.000%

En la siguiente gráfica se muestran los resultados de frecuencia de utilización de este cuadro.

FRECUENCIA UTILIZACION DAYFLO

Distribución de las 89 Consultas



Como se observa en los resultados obtenidos la base de datos más consultada es la del protocolo de los notarios, seguida de la Ley del Impuesto sobre la Renta y de las Doctrinas dictadas por los Tribunales, casi en un mismo porcentaje de utilización. La Constitución Política sigue en orden de utilización y por último están la Ley del Impuesto al Valor Agregado y la Ley de Papel Sellado y Timbres Fiscales.

Tomando en cuenta estos resultados y el lapso de tiempo en que se realizaron estas consultas podemos sacar las siguientes conclusiones sobre la frecuencia de uso de estos sistemas:

Tomando en cuenta que los tres meses tienen 4 semanas cada uno con 5 días hábiles cada una, llegamos a un total de 60 días hábiles como el lapso de esta prueba. Entonces, obtenemos los siguientes resultados en cuanto a la frecuencia diaria y semanal de utilización de los sistemas:

Tabla 6.3

Utilización por día
(60 días en total)

Base de Datos Consultada	Frecuencia	Consultas por Día
LEY IMPUESTO SOBRE LA RENTA	15	0.25
LEY IMPUESTO AL VALOR AGREGADO	8	0.133
CONSTITUCION POLITICA	10	0.166
LEY PAPEL SELLADO Y TIMBRES	8	0.133
GACETAS DE LOS TRIBUNALES	14	0.233
PROTOCOLOS	34	0.566
T O T A L	89	1.483

Para la utilización por semana se contaron todas las consultas en Dayflo a las diferentes bases de datos que se realizaron en cada una de las 12 semanas de duración de este estudio, obteniéndose la siguiente tabla de resultados:

Tabla 6.4

Utilización por Semana
(12 semanas en total)

No. de Semana	Numero de Consultas
1	10
2	7
3	7
4	9
5	8
6	6
7	5
8	7
9	6
10	7
11	8
12	9
T O T A L	89

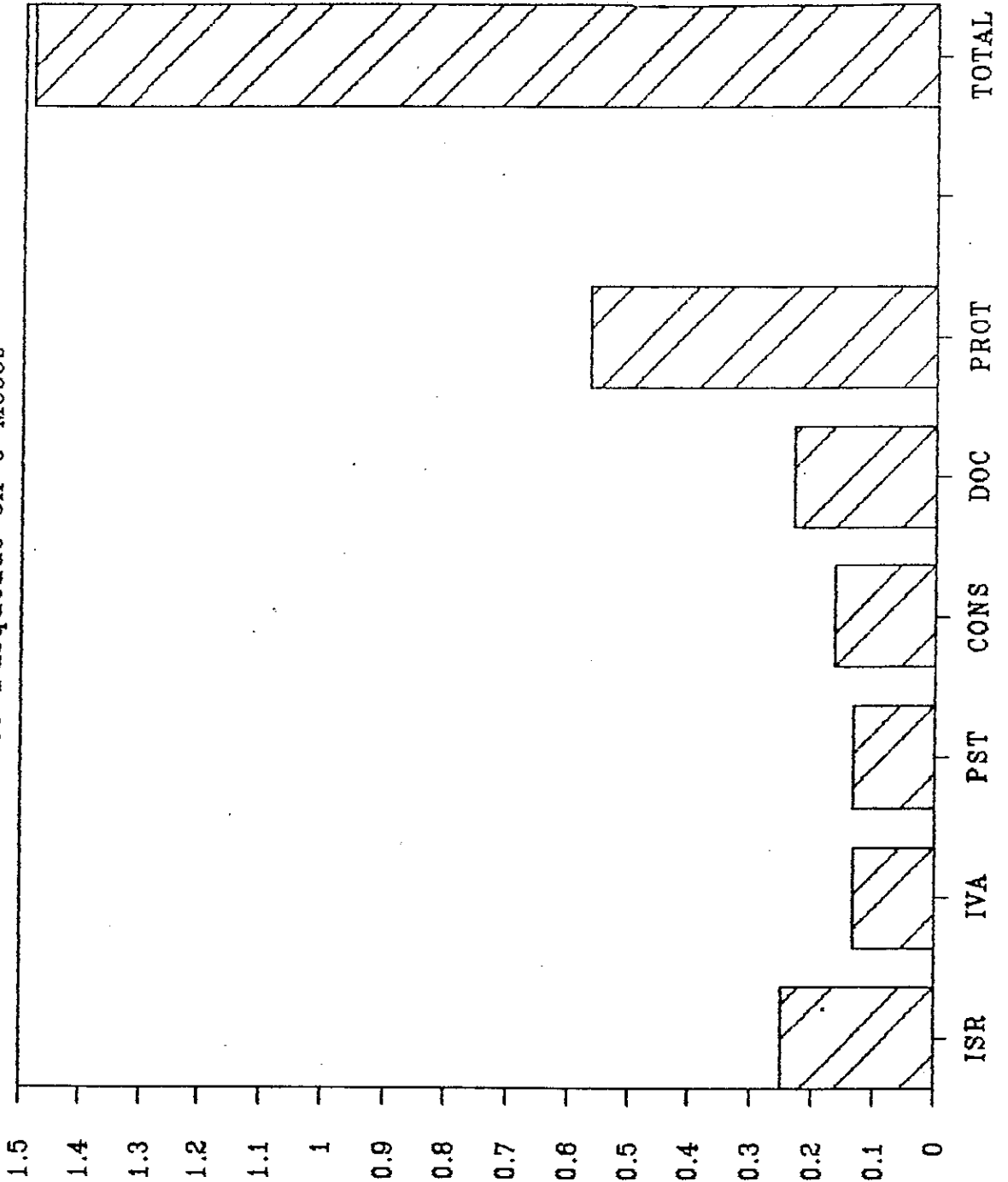
Media = 7.417 Desviación Típica = 1.382 Varianza = 1.910

Las siguientes dos gráficas muestran estos resultados:

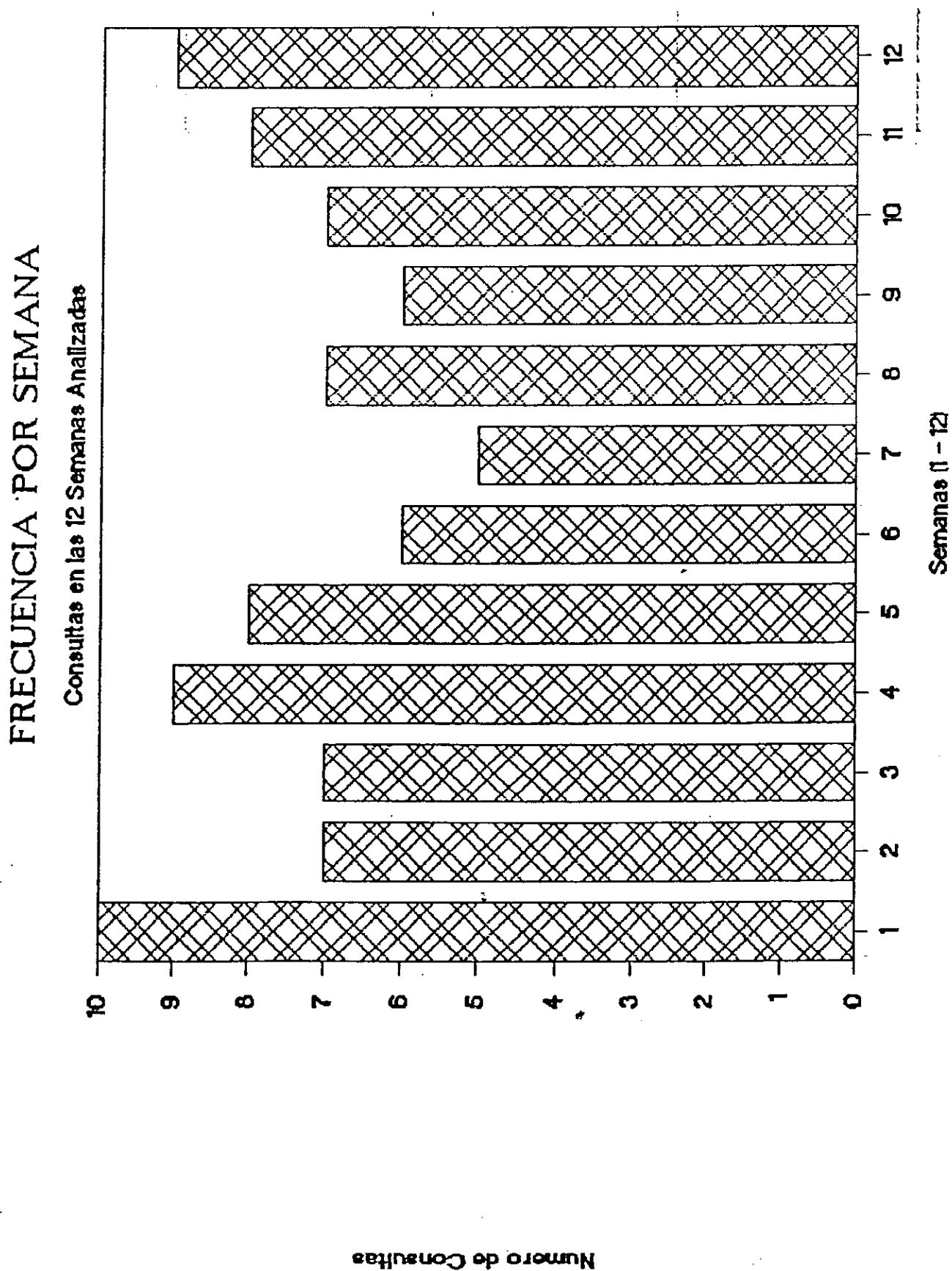
Gráfica 6.2

FRECUENCIA UTILIZACION POR DIA

89 Búsquedas en 3 Meses



Gráfica 6.3

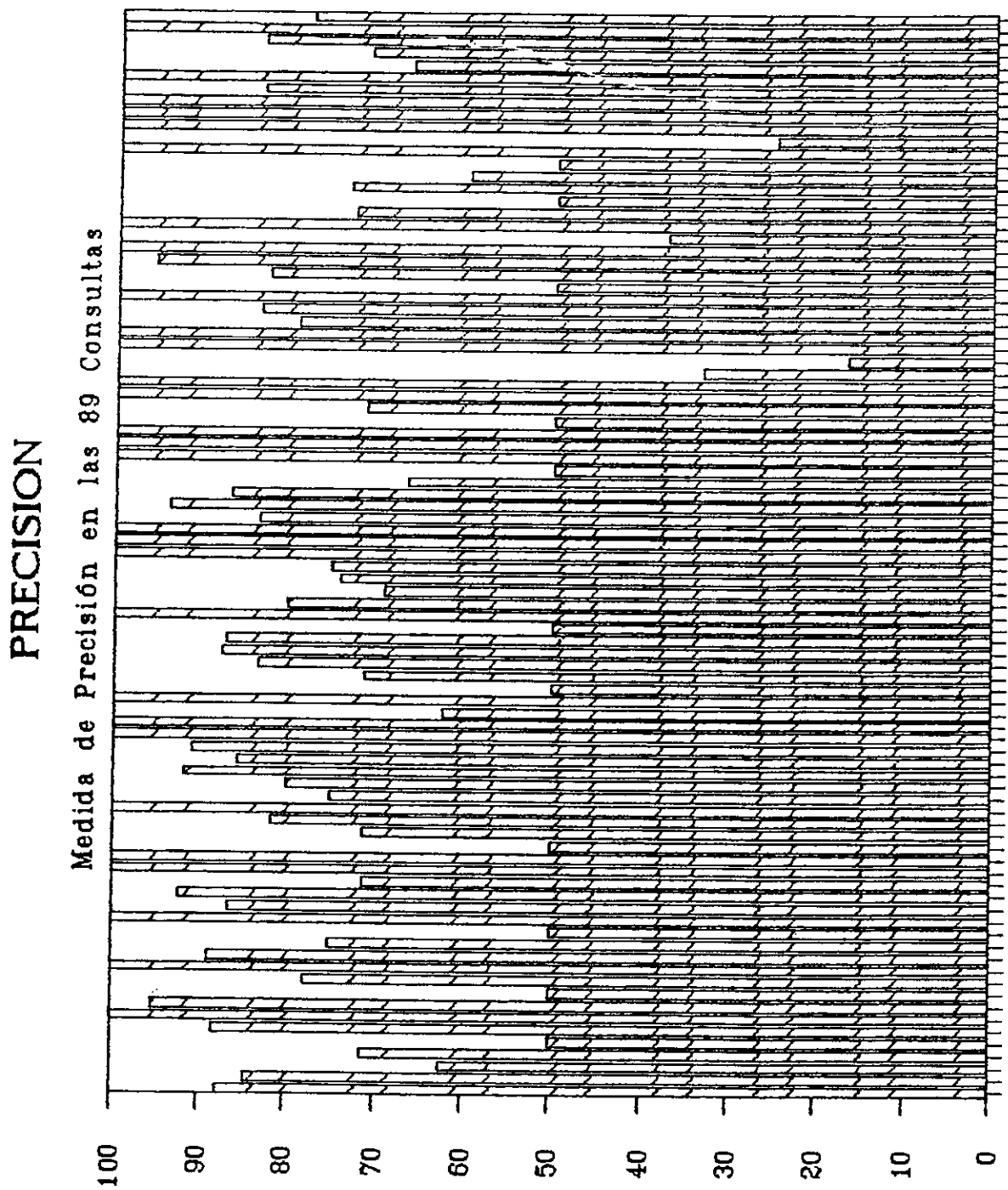


De acuerdo a estos resultados vemos que en total en promedio se hacen 1.483 consultas al día en este bufete y un total de 7.417 consultas a la semana. En los cuadros anteriores se puede observar la frecuencia diaria de utilización de cada base de datos independientemente. La frecuencia de uso de cada una de estas bases de datos depende del tipo de trabajos legales que se lleven a cabo en el bufete, pero los resultados obtenidos nos muestran que estos sistemas proporcionan una herramienta útil de trabajo.

b) Precisión.

La precisión de cada una de las consultas realizadas se calculó dividiendo el número total de documentos relevantes recuperados entre el total de documentos recuperados, en porcentaje. En el Anexo "D" se muestra el resultado obtenido en cada una de las consultas, variando este porcentaje de precisión entre un 16.667% y un 100.00%, con promedio de 80.247, una desviación típica de 20.184 y una varianza de 407.409. La varianza es el cuadrado de la desviación típica, y es la medida de dispersión o variación de los valores de la muestra alrededor de la media o promedio. En este caso obtuvimos una varianza bastante alta, por lo que deducimos que los valores de precisión están bastante dispersos con respecto al promedio. En la siguiente gráfica se muestran estos resultados.

Gráfica 6.4



Consultas (1 - 89)

Porcentaje de Precisión

c) Información adicional encontrada en las búsquedas.

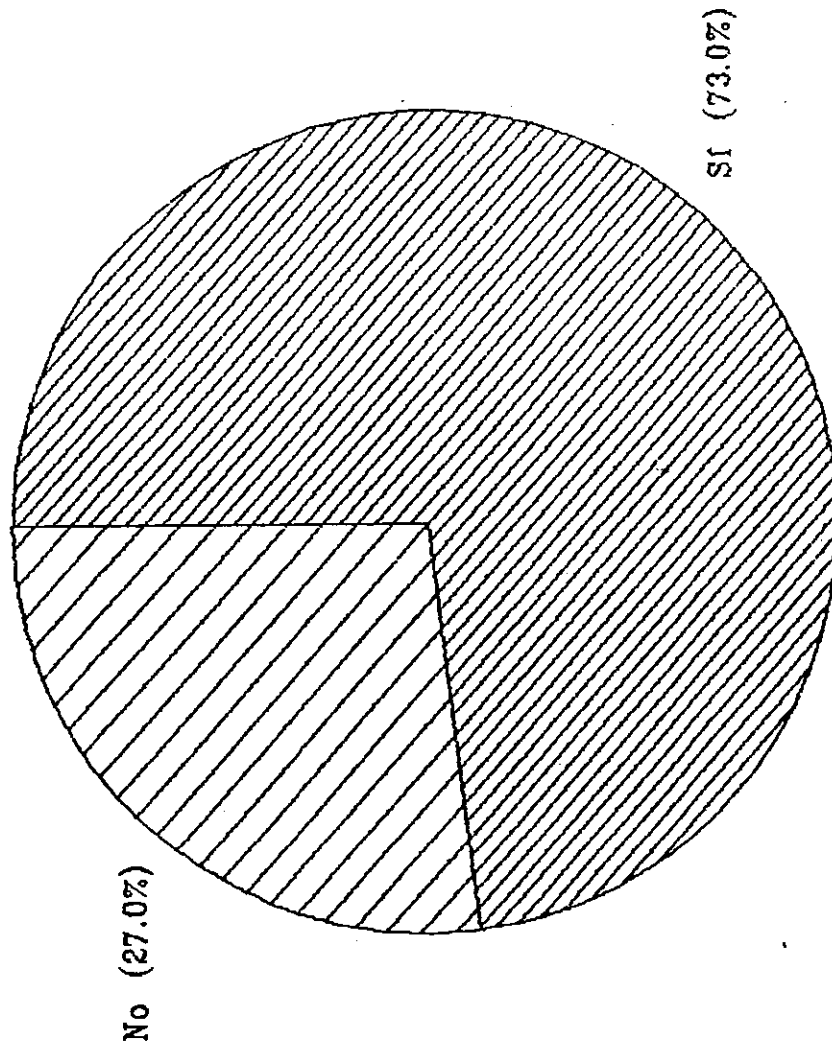
En la segunda pregunta del formulario de las consultas realizadas en esta prueba se anotó si se encontró en cada consulta información que no se había tomado en cuenta antes de hacer la búsqueda. Los resultados obtenidos se muestran también en el Anexo "D".

De las 89 consultas 65 afirmaron que se encontró información adicional al finalizar la búsqueda y 24 lo negaron. Esto nos da un 73.034% de resultados positivos y un 26.966% de resultados negativos. La mayor parte de los casos negativos se debió a que se quería encontrar un documento que ya se sabía que existía o que el abogado, según su experiencia y conocimientos, ya tenía en mente la información que podía esperar como resultado de las búsquedas.

En la siguiente gráfica se visualiza mejor este resultado obtenido.

Gráfica 6.5

INFORMACION ADICIONAL ?
Resultado de 89 Consultas



3. Análisis y conclusiones de los resultados obtenidos en la investigación de utilidad de este sistema.

En esta parte se presentan las conclusiones del análisis de utilidad desarrollado en esta sección.

Como se vió, el análisis cubrió la evaluación de tres puntos principales:

- Frecuencia de Utilización
- Precisión
- Utilidad para el profesional del derecho

En cuanto a la frecuencia de utilización de estos sistemas de bases de datos textuales en un bufete de abogados guatemalteco, se puede decir que tiene una buena frecuencia de uso, analizando los resultados obtenidos en esta evaluación. Se obtuvo un promedio de 1.483 consultas diarias a todas las bases de datos consideradas en este análisis.

Individualmente en cada base de datos también se midió la frecuencia de uso diario. El orden en que quedaron estas base de datos fué el siguiente:

- 1.- Protocolos,
- 2.- Ley del Impuesto sobre la Renta,
- 3.- Gacetas de los Tribunales,
- 4.- Constitución Política,
- 5.- Ley del Impuesto al Valor Agregado, y

6.- Ley del Papel Sellado y Timbres Fiscales.

Esta frecuencia de uso individual de cada base de datos está muy relacionada con el tipo de trabajo que se realice en el bufete de abogados y notarios. En este bufete los índices del protocolo de los notarios son los más utilizados. Esto se puede deber a que se realicen más trabajos notariales que de otra índole. A pesar de este hecho, toda la información considerada en este estudio fué consultada con una frecuencia aceptable.

Para el análisis de frecuencia de uso semanal se agruparon las consultas realizadas por cada una de las 12 semanas que duró el período de prueba para este análisis. El promedio semanal fué de 7.417 consultas, con una desviación típica de 1.382 y una varianza de 1.910. El dato de la varianza y de la desviación nos dan la pauta para determinar si los valores de los resultados de frecuencia de uso semanal estaban muy dispersos con respecto a la media. En nuestro caso el resultado fué que los datos se comportaron sin mayores variaciones con respecto a la media.

Al analizar el resultado obtenido con respecto a la precisión en las consultas realizadas, se puede establecer un porcentaje promedio de precisión de 80.743, el cual es bastante bueno. En este análisis de precisión se obtuvo una

desviación típica de 20.235 y una varianza de 409.470, lo que indica que los datos individuales estuvieron bastante dispersos con respecto a la media de los mismos.

El tipo de pregunta y la forma de transmitírselo al sistema de base de datos textual puede afectar la precisión y recolección en las búsquedas. En los casos en que se buscan raíces de palabras o palabras que puedan estar contenidas en otras palabras de significado diferente, se pueden recuperar un margen mayor de registros irrelevantes.

El tercer aspecto analizado fué el de la utilidad que un sistema de este tipo le puede proporcionar al profesional del derecho. Para medir este concepto se hizo la siguiente pregunta en cada una de las consultas realizadas: 2.- Encontró información que no había tomado en cuenta antes de hacer la búsqueda? Sí__ No__.

De las 89 consultas realizadas 64 contestaron afirmativamente esta pregunta y 25 negativamente, lo que nos da un porcentaje de 71.910% de casos afirmativos y un 28.090 de casos negativos.

Mucho que ver en este asunto lo tiene el tipo de pregunta que se le haga al sistema. Al observar la mayor incidencia de respuestas negativas se ve que se dan en

búsquedas realizadas en el Protocolo. Esto es muy interesante, ya que la mayoría de preguntas realizadas para el protocolo fueron de encontrar una escritura específica que ya se sabía de antemano que existía, con el propósito de localizarla rápidamente al encontrar su ubicación en el registro de la misma.

Por otro lado, las otras respuestas negativas se deben a que el abogado que hace la pregunta tiene mucha experiencia en su área y ya tiene el conocimiento de todo lo que puede esperar encontrar al realizar la búsqueda. Este punto es muy importante, ya que puede dar lugar a pensar que entonces la búsqueda en la base de datos textual es innecesaria en estos casos. Al surgir esta inquietud se les preguntó a los abogados que tuvieron el mayor número de consultas con este resultado su opinión al respecto. Ellos manifestaron que la información era muy útil y que proporcionaba una herramienta de trabajo muy valiosa, ya que los reportes generados por el sistema contemplaban toda esa información en un solo documento y daban información adicional en cuanto a fechas, ubicación o numeración según fuera el caso. Todos manifestaron que a pesar de la experiencia y conocimiento previo que un abogado o notario pueda tener de la información que necesita, todos somos seres humanos falibles y capaces de olvidar en cualquier momento alguna información vital al trabajo que se esté desarrollando. El sistema de bases de

datos textuales sirve entonces también como una confirmación a la información que se quiere utilizar en algún trabajo específico.

VII. SUGERENCIAS Y CONCLUSIONES

En este trabajo se cubrieron tres partes principales. La primera fue la parte teórica; la segunda fue el desarrollo de una aplicación con bases de datos textuales; y la tercera un análisis de utilidad de estos sistemas. En la primera parte se expuso la teoría que existe para estos sistemas de Bases de Datos Textuales o de formato libre. En la segunda parte se desarrolló una aplicación utilizando estos conceptos y sistemas innovadores. Y, por último, se hizo una evaluación de los mismos, con el propósito de tener un poco de métrica en el trabajo realizado.

La Aplicación que se desarrolló en este trabajo es pequeña comparando los alcances que puede tener un sistema de este tipo. Con un sistema de este tipo cualquier abogado o notario podría tener organizada toda su información textual de consulta de una manera desordenada pero fácil de encontrar en cualquier momento.

El sistema Dayflo no es el único ni el mejor que existe actualmente en el mercado de programas de software, pero si permite visualizar claramente el funcionamiento y capacidad de un sistema de base de datos de forma libre. El procesador de macros utilizado sirvió para poder hacer la aplicación específica de la Ley del Impuesto sobre la Renta. Este procesador de macros es bastante bueno aunque presenta algunas limitaciones en cuanto al tamaño de los macros y al manejo de situaciones condicionales.

Como conclusión, podemos decir que estos sistemas de Bases

de Datos Textuales son una herramienta muy poderosa para manejar cualquier tipo de información. Por ejemplo, en el caso específico del bufete de abogados y notarios, se pueden desarrollar un sin fin de aplicaciones como las que mencionamos anteriormente en las secciones B. "Descripción del ambiente de trabajo en un bufete profesional de abogacía y notariado" y A. "Tareas automatizables utilizando Bases de Datos Textuales" del capítulo IV de este trabajo. El tiempo es hoy en día para todo profesional un elemento vital para el éxito de sus proyectos. Si no se utilizan herramientas que faciliten y optimicen las tareas automatizables que consumen mucho tiempo en vano, el trabajo de estas personas se tornará poco eficiente.

El sistema de Base de Datos Textual Dayflo tiene mucha más flexibilidad que la presentada en esta aplicación. Se pueden realizar varias aplicaciones de varios tipos, ya sea combinando varias leyes, códigos, decretos o resoluciones importantes y realizar búsquedas sobre esta información. El índice de Protocolo del notario es otra aplicación que se puede desarrollar en un sistema de este tipo, pudiéndose encontrar en cualquier momento cualquier información que esté contemplada en el mismo, así como también imprimir reportes y hacer estadísticas de esta información.

Con un sistema de este tipo el trabajo de buscar información textual por horas se reduce a pocos minutos o segundos

dependiendo del volumen de información consultada. Realmente es impresionante cómo se logra encontrar cualquier material textual en un sistema de Base de Datos Textual. El espacio de almacenamiento de la información es mucho más reducido que las tradicionales estanterías para almacenar papeles, ya que todo se guarda en medios magnéticos. Conforme avance la tecnología estas ventajas serán cada vez más eminentes.

Uno de los puntos más discutidos con estos sistemas de consulta de información textual es si se debe utilizar un sistema que almacene un resumen del texto general ó todo el texto para los efectos de las búsquedas de información. Realmente dependiendo del tipo de información que se quiera almacenar así se puede utilizar el uno o el otro. Específicamente, para lo que son leyes el sistema de base de datos debe ser, definitivamente, de búsqueda completa sobre todo el texto de la ley. En base a la experiencia y en base a pláticas sostenidas con varios profesionales del Derecho en Guatemala, pude establecer que con las leyes no sirve buscar información sobre la interpretación que otra persona le haya podido dar a la misma, entonces si se utiliza un sistema de índices contruídos en base a una interpretación humana del texto de la ley, se puede caer en un margen grande de desacierto en las búsquedas. Por lo tanto se debe utilizar un sistema de búsqueda en textos completos que sea lo suficientemente rápido y confiable.

Para otras aplicaciones o para la de consulta de leyes realmente lo ideal es utilizar un sistema de base de datos que tenga los dos sistemas de búsqueda de la información combinados para que así el sistema sea realmente eficiente.

Para escoger qué Base de Datos y que computador se debe utilizar para instalar un sistema de consulta de información textual, como el desarrollado en este trabajo, se debe también tomar en cuenta los siguientes parámetros:

- 1- la cantidad de información que se quiere almacenar en el sistema;
- 2- el tipo de consultas que se querrán realizar sobre la información del sistema;
- 3- la forma de la información que se quiere almacenar en el sistema;
- 4- la velocidad de respuesta para las consultas que se realicen;
- 5- el tiempo de acceso a la información;
- 6- si es importante marcar la posición exacta de las palabras buscadas dentro de los textos encontrados;
- 7- la flexibilidad que exista en cuanto a la definición de nuevos elementos de texto;
- 8- la habilidad de manejar conexiones con documentos de otros programas o aplicaciones externas como procesadores de palabras, hojas electrónicas, bases de datos, etc.; y

- 9- si la Base de Datos Textual tiene otros métodos de búsqueda de la información aparte de el de encontrar palabras en todo el texto.

El sistema Dayflo no tiene algunas de estas características, pero sí cumplió con los requisitos que se necesitaban para desarrollar la aplicación. Actualmente existen sistemas de Bases de Datos Textuales que tienen todas las características mencionadas arriba, o por lo menos la combinación de varias, como el sistema Dayflo.

La recomendación es analizar cuáles de las nueve características mencionadas son vitales para el funcionamiento de la aplicación que se quiere ejecutar y de esta forma buscar un sistema de Base de Datos Textual que llene los requisitos de la aplicación.

El análisis de utilidad realizado en el Capítulo VI de este trabajo nos proporcionó resultados muy interesantes.

La frecuencia de uso de estos sistemas en un bufete de abogados y notarios fue bastante buena. La media diaria de utilización fue de 1.483 consultas. También se vió que dependiendo del tipo de trabajo que se realice en el bufete, así es la frecuencia de consultas a una base de datos específica. En este análisis la base de datos de los índices de protocolo fue la

más consultada.

Con respecto a la precisión medida en este mismo análisis, se obtuvo una media de 80.743%. Los datos de la muestra, sin embargo, estuvieron bastante dispersos con relación a esta media, ya que la varianza fue alta (409.47).

En cuanto a la información adicional obtenida de las búsquedas realizadas, aproximadamente, el 72% de los usuarios del sistema respondieron que sí habían encontrado información adicional a la esperada, y un 28% respondió negativamente esta misma pregunta. Los casos negativos fueron, en su mayoría, debido a que el usuario sabía de antemano lo que quería encontrar, pero lo que necesitaba era información más detallada de los documentos a encontrar. En los índices de protocolo se dió mucho este tipo de situación.

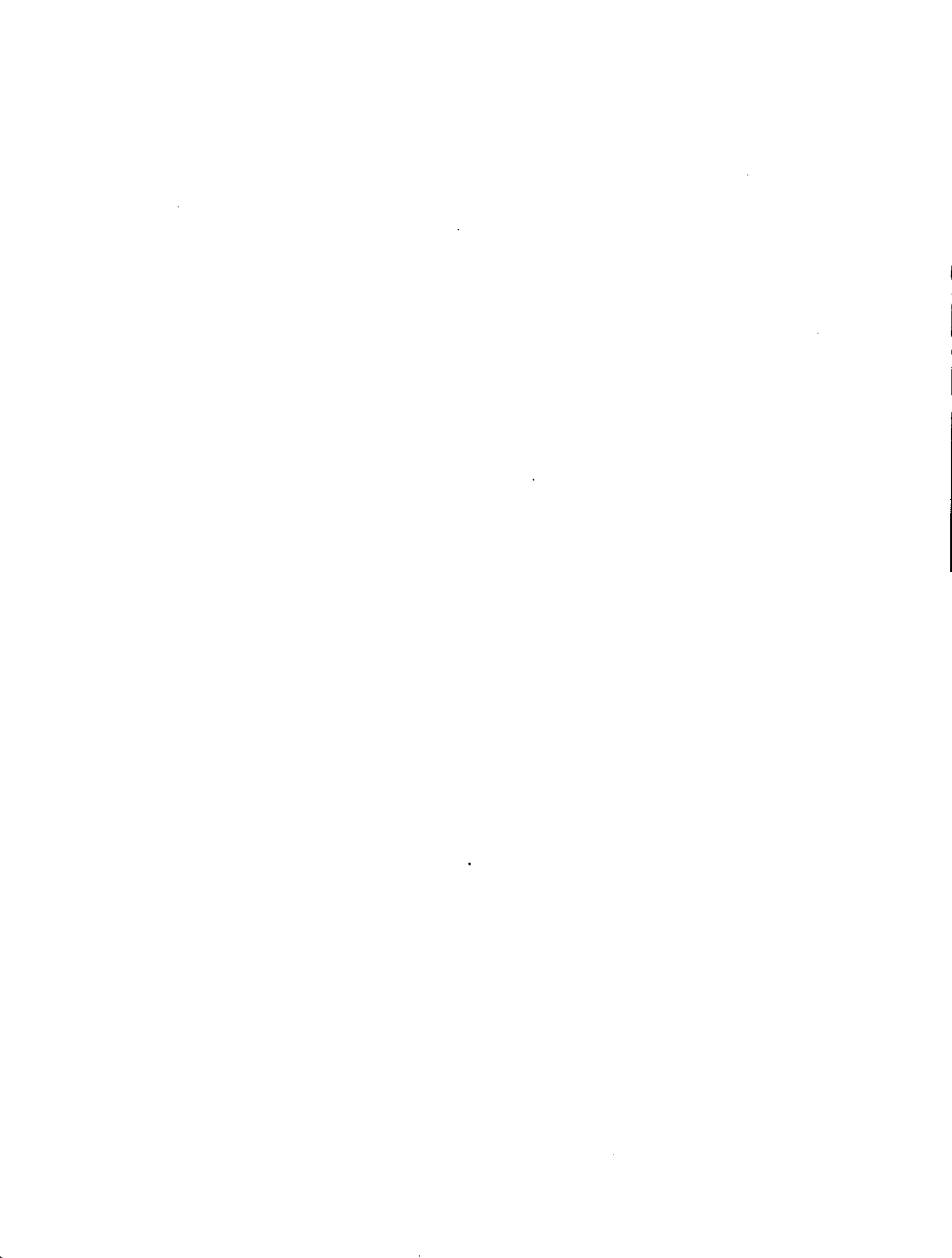
Otra causa de las respuestas negativas fue el conocimiento o experiencia que tienen los usuarios sobre la información consultada. Este tipo de situación me hizo pensar que entonces estos sistemas podrían no ser útiles en realidad, pero al interrogar al respecto a estos usuarios expertos, me manifestaron que, al contrario de lo supuesto, este sistema les era de mucha utilidad, tanto como una confirmación de sus conocimientos como de agilidad en tener toda esa información agrupada en corto tiempo.

VIII. BIBLIOGRAFIA

- Aguirre Godoy, Mario.** "Repertorio de Jurisprudencia". Tomo I y 1962 II. Editorial Universitaria. Guatemala.
- Altman, Mary Ann and Weil, Robert.** "How to Manage your law 1987 office". Matthew Bender & Co.Inc. New York.
- Altmark, Daniel R.** "Informática y Derecho". Editorial Depalma, 1987 Buenos Aires, Argentina.
- Arentowicz, Frank and Bower, Ward.** "Law office Automation and 1987 Technology". Matthew Bender & Co.Inc. New York.
- Begeman, Michael and Conklin Jeff.** "The Right Tool for the Job". 1988 BYTE the Small Systems Journal, McGraw-Hill Publication, Octubre. 255.
- Blair, David and Maron, M.E.** "An Evaluation of Retrieval 1985 Effectiveness for a Full-Text Document-Retrieval System". Communications of the ACM, Vol. 28, No. 3, March.
- Burke, William J. and Bradbury, Carl W.** "Accounting Systems for 1988 Law Offices". Matthew Bender & Co. Inc., N.Y.
- Centro Regional del IBI para la enseñanza de la Informática.** 1984 "Actas: I Congreso Iberoamericano de Informática Jurídica". Santo Domingo.
- Date, C.J.** "An Introduction to Data Base Systems". Vol.1, 4a. 1987 Edición, Addison-Wesley Systems Programming Series, California.
- Faloutsos, Christos.** "Access Methods for Text". Computing 1985 Surveys, Vol.17, No.1, March.
- Fiderio Janet.** "A Grand Vision". Byte The Small Systems Journal, 1988 McGraw-Hill Publication, Octubre.
- Frisse Mark.** "From Text to Hypertext". Byte The Small Systems 1988 Journal, McGraw-Hill Publication, Octubre.
- Guatemala. Organismo Judicial. Gacetas de los Tribunales.**
---- Guatemala.
- Hildebrandt, Bradford W.** "Law Office Management 1983". 1983 Practising Law Institute.
- Jordán Flórez, Fernando.** "La Informática, el Estado y el 1984 Derecho". Universidad Piloto de Colombia (CII-UP). Bogotá.

- _____. "La Informática Jurídica". Universidad Piloto de
1985 Colombia (CII-UP). 2a. Edición. Bogotá. 212 pp.
- Kubernan: Legal Management System.** Descripción Operacional.
1980 Winston-Salem. s/n.
- Landsburg, Harry.** "Planning for Computers. Evaluating Data
1981 Processing Needs for Medium and Large Law
Firms". Section of Economics of Law Practice, American
Bar Association. Illinois.
- López-Muñiz Goñi, Miguel.** "Informática Jurídica Documental".
1984 Díaz de Santos, S.A. 1a. Edición. Madrid.
- Martin, James.** "Principles of Data-Base Management". Prentice
1976 Hall Inc.. New Jersey.
- Salton Gerald.** "Another look at Automatic Text-Retrieval
1986 Systems". Communication of the ACM, Vol.29, No.7.
- Shapiro Ezra.** "Text Databases". Byte The Small Systems Journal,
1984 Vol.9, No.11, McGraw-Hill Publication.
- Toffler, Alvin.** "La Tercera Ola". Plaza & Jones Editores, S.A.,
1984 Barcelona, España.
- Ullman, Jeffrey D.** "Principles of Database Systems". Computer
1980 Science Press Inc.. Maryland. 379 pp.
- Wiederhold, Gio.** "Database Design". McGraw-Hill Inc.. New York.
1977 658 pp.
- Withgott, Coleen and Anderson, Austin.** "Webster's legal
1981 secretaries Handbook". G&C Merriam Company,
Publishers. Massachusetts.

"ANEXOS"



Anexo "A"

PROGRAMAS UTILIZADOS PARA DESARROLLAR LA APLICACION

```
echo off
cls
smartkey >nul:
skbatch /L g:\Dayflo\LEYISR2.dtx >nul:
Dayfloex g:\Dayflo
skbatch /u
copy g:\Dayflo\adios.msg con: >nul:
```

Key	Definition
<[SS A]>	<<W2,2>[Alt H]16[Alt W]40[PgUp][PgDn][Alt T]Menu de Mant enimiento[Enter][Enter] [F2] SISTEMA DE CONSULTAS [F1][Enter] LEY DEL IMPUESTO SOBRE LA RENTA[Ent er][Enter][Enter] [F9][F8][F1][F1] Revisar/Modificar A rticulos[Enter][Enter] [F8][F2][F1] Ordenar la Ley[Ent er][Enter] [F8][F3][F1] [F10]Regreso a Menu anterior[E nter][Enter][Enter][Enter][Enter][Enter]<M[Alt M]> <*>Mantenimiento de la Informacion>
<[SS C]>	<<E[SS G]><E[SS H]><E[SS I]><E[SS F]>>
<[SS D]>	<[Alt F2]LULEY[Enter][^]><E[SS w]>>
<[SS E]>	<<E[SS W]>[Alt F2]LULEY[Enter][Alt F2]LUORDENAR[Enter][A lt F2]LUBLANCO[Enter]<E[SS A]>>
<[SS F]>	<<E[SS P]>>
<[SS G]>	<[F3]N20[Enter][F3]RS1[Enter]TEXTO[Enter][^]C[Enter]<E[SS N]>>
<[SS H]>	<[^][F3]N20[Enter][F3]RS1[Enter]TEXTO[Enter][^]C<E[SS u]><P>[Enter][^][Enter][Enter]TEXTO[Enter][^]C<E[SS U ><P>[Enter][Enter]R[Enter]<E[SS n]>>
<[SS I]>	<[F3]N19[Enter][^][Alt F3]GINDICE[Enter]<E[SS Z]>>
<[SS J]>	<[^]><E[SS l]>>
<[SS K]>	<<E[SS k]>[^][Alt F4]R[F8][Enter]PLEYISR_GRAL[Enter]PQ< E[SS n]>>
<[SS L]>	<[F3][F8][Enter]M[Enter]S1[Enter][^][F3]N11[Enter][^]< E[SS b]>>
<[SS N]>	<<W15,30>[Alt H]05[Alt W]20[PgUp][PgDn][Alt T]Frases a b uscar[Enter][Enter] Ingrese las frases[Enter] que quiere buscar.[Enter][Enter] [F4]Alt-C[F1] AYUDA<K0> <*>PANTALLA INGRESO PALABRAS>
<[SS P]>	<<W2,2>[Alt H]16[Alt W]40[PgUp][PgDn][Alt T]Menu Princip al[Enter][Enter] [F2] SISTEMA DE CONSULTAS [F 1][Enter] LEY DEL IMPUESTO SOBRE LA RENTA[Enter][Ent er][Enter] [F9][F8][F1][F1] Mantenimiento de la Inform acion[Enter][Enter] [F8][F2][F1] Consultas y Reportes[Enter][Enter] [F8][F3][F1] [F10]Salida del Sistema[Ent er][Enter][Enter][Enter][Enter][Enter]<M[SS S]>>
<[SS S]>	<<E[SS A]><E[SS b]><E[SS c]>>
<[SS U]>	<<W10,20>[Alt H]05[Alt W]40[PgUp][PgDn][Alt T]INGRESE LA SEGUNDA FRASE[Enter][Enter] [F4]Ingrese la segunda fra se a buscar.[F1][Enter][Enter] <Presione cualquier tecla para seguir>[Enter]<K0> <*>segunda frase>
<[SS W]>	<<W10,20>[Alt H]10[Alt W]40[PgUp][PgDn][Alt T]Proceso de Ordenamiento[Enter][Enter] El sistema procedera a ord enar la[Enter] Ley por No. de Artículo e Inicisos.[Ent er][Enter][Enter] [F9][F4]Favor esperar hasta regresar al MENU.[F1][Enter][Enter] G R A C I A S[Ent er][Enter] <pulse cualquier tecla><K0> <*>PANTALLA ORDENAR>
<[SS X]>	<<A>D[Alt Z]R<P> <*>GENERA REPETICION DE BUSQUEDAS AL INDICE DE PALABRAS>
<[SS Z]>	<<W2,39>[Alt H]07[Alt W]40[PgUp][PgDn][Alt T]INGRESO LIS TA DE PALABRAS[Enter][F9][Enter][Enter][F10] [F9][F4][Al t]-O[F1] Para ingresar otra palabra[Enter][Enter] [F4][Alt]-T[F1] Cuando Termine[Enter][Enter]<K0> <*>INGRESO LISTA PALABRAS INDICE>

Key	Definition
<[SS b]>	<<W2,2>[Alt H]16[Alt W]40[PgUp][PgDn][Alt T]Menu de Consultas y Reportes[Enter][Enter] [F2] SISTEMA DE CONSULTAS [F1][Enter] LEY DEL IMPUESTO SOBRE LA RENTA[Enter][Enter][Enter] [F9][F8][F1][F1] Articulos que e contengan 1 o va- [Enter] rias palabras o frases (Max.10).[Enter] [F3][F2][F1] Articulos que contengan 2 pala- [Enter] [F10]bras o frases en el mismo.[Enter][Enter] [F9][F3][F3][F1] Generar Indice de una lista de [Enter] palabras.[Enter][Enter] [F3][F4][F1] Regreso a Menu anterior[Enter]<M[SS C]> <*>Consultas y Reportes
<[SS c]>	<[F3]n12[Enter][F3]c[~]]q>
<[SS j]>	<<E[SS J]><E[SS K]><E[SS L]>>
<[SS k]>	<<W15,10>[Alt H]05[Alt W]60[PgUp][PgDn][Alt T]ALISTE LA IMPRESORA[Enter][Enter][Enter] Presione cualquier tecla cuando [Enter] la impresora este lista.[Enter][Enter]<K0> <*>pantalla imprimir
<[SS l]>	<<W2,20>[Alt H]16[Alt W]40[PgUp][PgDn][Alt T]REVISION EN PANTALLA[Enter][F9][Enter][Enter][F10] [F9][F4][F10][F1] Avanzar Registro[Enter][Enter] [F4][F9][F1] Regresar Registro[Enter][Enter] [F4][^X] [^Y][F1] Linea arriba y abajo respectiva- mente[Enter][Enter] [F4][^][^Z][F1] Caracter a la izquierda y derecha respectivamente[Enter][Enter][Enter] [F4][Alt]-N[F1] CUANDO TERMINE[Enter][Enter]<K0> <*>REVISION EN PANTALLA
<[SS n]>	<<W2,10>[Alt H]16[Alt W]40[PgUp][PgDn][Alt T]Resultado de la Consulta[Enter][Enter][Enter][Enter][Enter][Enter] [F9][F4][F1][F1] Ver Articulos en Pantalla[Enter][Enter] [F4][F2][F1] Imprimir Articulos encontrados[Enter][Enter] [F4][F3][F1] Regresar a Menu Anterior[Enter][Enter][Enter][Enter][Enter]<M[SS j]> <*>RESPUESTA DE CONSULTAS
<[SS r]>	<[K*]cv[Enter][F3]n20[Enter][F3]rsl[Enter]texto[Enter]c[~P][Enter][Enter]r[Enter][Alt F4]r[F8][Enter]pindice[Enter]findice.txt[Enter]yq[F3]mc[F8][Enter]sl[Enter][Alt F2]tfindice.txt[Enter][F7][^PgDn][^Home][^End][Left][K*]c[Enter][F3][F3]mct[F3]n19[Enter][K*]pd[F10]>
<[SS t]>	<<E[SS k]>[Alt F4]r[F8][Enter]pindice final[Enter]pq[F3]m[F8][Enter]ct[F3]n11[Enter][^]][^]<E[SS b]>>
<[SS u]>	<<W10,20>[Alt H]05[Alt W]40[PgUp][PgDn][Alt T]INGRESE LA PRIMERA FRASE[Enter][Enter] [F4]Ingrese la primera frase a buscar.[F1][Enter][Enter] <Presione cualquier tecla para seguir>[Enter]<K0> <*>primera frase
<[SS v]>	<<W2,20>[Alt H]16[Alt W]40[PgUp][PgDn][Alt T]PRESIONE LAS SIGUIENTES TECLAS:[Enter] [Ctrl]-[F5][Enter] C[Enter] A[Enter] R[Enter] No. Palabras de la lista[Enter][Enter][Enter] [Ctrl]-[F5][Enter] C[Enter] E[Enter] [F5]-r[Enter] [Ctrl]-[F5][Enter] C[Enter] E[Enter] [F5]-t[Enter] [Ctrl]-[F5][Enter] [Enter][Enter] [Alt]-Z<K0> <*>NUMERO DE PALABRAS
<[SS w]>	<<W20,59>[Alt H]04[Alt W]20[PgUp][PgDn][Alt T]Edicion Ar

Key	Definition
	-iculos[Enter][Enter] AYUDA ... [Alt]-A[Enter][Enter]Enter para seguir. <K0>>
<[Alt E]>	<[F3]nll[Enter][^[]<E[SS A]>>
<[Alt R]>	<<A>D[Alt 3]Z>
<[Alt T]>	<[F8][Enter][F9]<E[SS v]><E[SS X]>>
<[Alt U]>	<[^[] [F3]nll[Enter][^[] [^[]<E[Alt H]>>
<[Alt I]>	<[Alt F3]gleyisr[Enter]>
<[Alt O]>	<[F10][Alt F3]GINDICE[Enter]>
<[Alt A]>	<<W2,49>[Alt H]16[Alt W]30[PgUp][PgDn][Alt T]AYUDA EDITOR[Enter][Enter] [F9][F4][Alt]-A[F1] AYUDA[Enter][Enter] [F10][F4][F9][F1] Regresar Registro[Enter] [F4][F10][F1] Avanzar Registro[Enter][Enter] [F4][F6][F1] Avanzar Campo[Enter] [F4][Shift]-[F6][F1] Regresar Campo[Enter][Enter] [F4][Alt]-I[F1] Agregar Registro[Enter] [F4][Alt]-B[F1] Borrar Registro[Enter][Enter] [F4][Ctrl]-[End][F1] Parrafo sig.[Enter] [F4][Ctrl]-[Home][F1] Parrafo ant.[Enter][Enter] [F9][F4][Alt]-E[F1] SALIDA<K0>>
	<*>AYUDA EDITOR>
<[Alt S]>	<[Enter]r[Enter]<E[SS n]>>
<[Alt H]>	<<W2,2>[Alt H]16[Alt W]40[PgUp][PgDn][Enter][Enter] S T E M A D E C O N S U L T A[Enter][Enter][Enter][Enter] [F3] [F1][Enter] [F3] LEY DEL IMPUESTO SOBRE LA RENTA [F1][Enter] [F3] [F1][Enter][Enter][Enter][Enter]Presione cualquier tecla para continuar<K0><E[SS P]>>
	<*>PRESENTACION>
<[Alt Z]>	<<R3><E[SS r]><E[SS t]>>
<[Alt C]>	<<W1,20>[Alt H]16[Alt W]40[PgUp][PgDn][Alt T]AYUDA INGRESO DE FRASES A BUSCAR[Enter] [F9][F4]Alt-C[F10][F1] [F2]AYUDA[F1][Enter] Ingrese las frases o palabras[Enter] separadas por Enter. Los siguientes[Enter] simbolos pueden usarse para expresiones completas:[Enter] [F9][F4]*[F1] representa un cadena de cualquier[Enter] caracter excepto espacio y de longitud variable.[Enter] [F4]@[F1] representa cualquier letra sola.[Enter] [F4]#[F1] representa cualquier digito solo.[Enter] [F4]_[F1] representa cualquier caracter excepto espacio.[Enter] [F4]?[F1] indica que el caracter anterior es opcional.[Enter] [F10][F2]Maximo de frases = 10[F9][F1][Enter] Al ingresar la ultima: [F4][Alt]-S[F1]<K0>>
	<*>AYUDA INGRESO PALABRAS>
<[Alt B]>	<[F3]M[Enter]T[Enter][Up]>
<[Alt N]>	<[^[]<E[SS n]>>
<[Alt M]>	<<E[SS D]><E[SS E]><E[SS F]>>
<[F6]>	<[^PgDn]>
<[Sh F6]>	<[^PgUp]>



Anexo "B"

PANTALLA DE DEFINICION DE REGISTROS PARA LA APLICACION

A LEY DEL IMPUESTO SOBRE LA RENTA

Ley IMPUESTO SOBRE LA RENTA

Capitulo I.- OBJETO Y CAMPO DE APLICACION

Articulo 1 Epigrafe Objeto
Inciso Sub-Inciso

Texto El impuesto sobre la Renta es anual y recae sobre las rentas o utilidades que provengan de la inversión de capital, del trabajo o de la combinación de ambos y sobre los incrementos de patrimonio neto, que se produzcan en el período anual de imposición, generado por rentas gravadas no declaradas, más, en su caso, los gastos establecidos en esta ley y su reglamento.

ayFlo: Press F1 for help



Anexo "C"

ELEMENTOS DEFINIDOS EN EL DICCIONARIO DE LA BASE DE DATOS
TEXTUAL PARA EL DESARROLLO DE LA APLICACION DE LA LEY DEL
IMPUESTO SOBRE LA RENTA.

Dictionary for c:\dayflo\isr1.db

Item Name	No.	Indx	Interp	Length	Pattern
Articulo	4	y	Number	4	
Capitulo	6	y	Text		
Epigrafe	5	y	Text		
Inciso	1002	n	Text	5	
Ley IMPUESTO SOBRE LA RENTA	3	y	Text		
Sub-Inciso	1003	n	Text	5	
Texto	7	y	Text		



ANEXO "D"

RESULTADOS OBTENIDOS DE LAS CONSULTAS REALIZADAS EN UN BUFETE DE ABOGADOS DURANTE TRES MESES

No. de Busqueda	Base de Datos Consultada	Registros		Precision	Informacion Util no prevista	
		Encontrados No.	Relevantes No.		Sí	No
1	CONSTITUCION	49	43	87.755	x	
2	CONSTITUCION	39	33	84.615	x	
3	CONSTITUCION	8	5	62.500	x	
4	CONSTITUCION	7	5	71.429	x	
5	CONSTITUCION	2	1	50.000		x
6	CONSTITUCION	17	15	88.235	x	
7	CONSTITUCION	4	4	100.000	x	
8	CONSTITUCION	21	20	95.238	x	
9	CONSTITUCION	2	1	50.000		x
10	CONSTITUCION	9	7	77.778	x	
11	DOCTRINAS	1	1	100.000	x	
12	DOCTRINAS	9	8	88.889	x	
13	DOCTRINAS	4	3	75.000	x	
14	DOCTRINAS	2	1	50.000		x
15	DOCTRINAS	3	3	100.000	x	
16	DOCTRINAS	15	13	86.667	x	
17	DOCTRINAS	13	12	92.308	x	
18	DOCTRINAS	28	20	71.429	x	
19	DOCTRINAS	1	1	100.000		x
20	DOCTRINAS	4	4	100.000	x	
21	DOCTRINAS	8	4	50.000		x
22	DOCTRINAS	49	35	71.429	x	
23	DOCTRINAS	22	18	81.818	x	
24	DOCTRINAS	3	3	100.000	x	
25	ISR	12	9	75.000	x	
26	ISR	10	8	80.000	x	
27	ISR	37	34	91.892	x	
28	ISR	14	12	85.714	x	
29	ISR	11	10	90.909	x	
30	ISR	4	4	100.000		x
31	ISR	9	9	100.000	x	
32	ISR	8	5	62.500	x	
33	ISR	6	6	100.000	x	
34	ISR	8	4	50.000	x	
35	ISR	14	10	71.429	x	
36	ISR	6	5	83.333	x	
37	ISR	56	49	87.500	x	
38	ISR	31	27	87.097	x	
39	ISR	2	1	50.000	x	
40	IVA	1	1	100.000		x
41	IVA	5	4	80.000	x	
42	IVA	13	9	69.231	x	
43	IVA	23	17	73.913	x	
44	IVA	8	6	75.000	x	
45	IVA	8	8	100.000	x	
46	IVA	17	17	100.000	x	
47	IVA	2	2	100.000	x	

RESULTADOS OBTENIDOS DE LAS CONSULTAS REALIZADAS EN UN BUFETE DE ABOGADOS DURANTE TRES MESES

No. de Busqueda	Base de Datos Consultada	Registros Encontrados	Registros Relevantes	Precision	Informacion Util no prevista	
		No.	No.		Si	No
48	PROTOCOLO	18	15	83.333	x	
49	PROTOCOLO	32	30	93.750	x	
50	PROTOCOLO	45	39	86.667	x	
51	PROTOCOLO	3	2	66.667	x	
52	PROTOCOLO	6	3	50.000		x
53	PROTOCOLO	1	1	100.000		x
54	PROTOCOLO	7	7	100.000	x	
55	PROTOCOLO	2	2	100.000	x	
56	PROTOCOLO	4	2	50.000		x
57	PROTOCOLO	14	10	71.429	x	
58	PROTOCOLO	6	6	100.000	x	
59	PROTOCOLO	18	18	100.000	x	
60	PROTOCOLO	3	1	33.333		x
61	PROTOCOLO	6	1	16.667		x
62	PROTOCOLO	2	2	100.000		x
63	PROTOCOLO	4	4	100.000	x	
64	PROTOCOLO	38	30	78.947		x
65	PROTOCOLO	6	5	83.333	x	
66	PROTOCOLO	1	1	100.000		x
67	PROTOCOLO	2	1	50.000		x
68	PROTOCOLO	57	47	82.456	x	
69	PROTOCOLO	23	22	95.652	x	
70	PROTOCOLO	2	2	100.000	x	
71	PROTOCOLO	8	3	37.500		x
72	PROTOCOLO	1	1	100.000		x
73	PROTOCOLO	33	24	72.727	x	
74	PROTOCOLO	2	1	50.000		x
75	PROTOCOLO	15	11	73.333	x	
76	PROTOCOLO	5	3	60.000	x	
77	PROTOCOLO	4	2	50.000		x
78	PROTOCOLO	1	1	100.000		x
79	PROTOCOLO	4	1	25.000		x
80	PROTOCOLO	1	1	100.000		x
81	PROTOCOLO	5	5	100.000	x	
82	PS Y TF	6	6	100.000	x	
83	PS Y TF	12	10	83.333	x	
84	PS Y TF	1	1	100.000		x
85	PS Y TF	3	2	66.667	x	
86	PS Y TF	7	5	71.429	x	
87	PS Y TF	12	10	83.333	x	
88	PS Y TF	2	2	100.000	x	
89	PS Y TF	9	7	77.778	x	

Media	80.247	65	24
Minimo	16.667	73.034 %	26.966 %
Maximo	100.000		
Desviacion	20.184		
Varianza	407.409		

ANEXO "E"
SOLICITUD DE INFORMACION

No.

Fecha: _____	Solicitante: _____
No. Expediente: _____	
Dónde desea buscar? _____ _____ _____	
Qué frases o palabras quiere buscar? _____ _____ _____	
Las frases deben estar en el mismo registro simultáneamente? Sí _____ No _____	
Con que una de las frases esté presente en el registro es suficiente para recuperar el registro? Sí _____ No _____	
Registros encontrados:	Registros buscados:

Fecha de entrega: _____

Firma de Recibido: _____

Favor responder las siguientes preguntas:

- 1.- De los registros encontrados por el sistema cuántos eran relevantes a la pregunta realizada? _____
- 2.- Encontró información que no había tomado en cuenta antes de hacer la búsqueda? Sí _____ No _____



ANEXO "F"

Arquitectura del Sistema DayFlo

